

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 7, 2002, 08:58:48 ; Search time 16678.2 Seconds
(without alignments)
12077.973 Million cell updates/sec

Title: US-09-150-867-2
Perfect score: 9626
Sequence: 1 gaattcgcgagtcgatatag.....ttaaaaaaaaaaacggaattc 9626

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sta:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_jmu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sta:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result Query #
No. Score Match Length DB ID Description

1	9610	99.8	9610	5	AF027728	AF027728 Xenopus 1
2	620.4	6.4	8257	9	HSCNPE	X21505 H.sapiens C
3	243.2	2.5	2576	8	AB028468	AB028468 Arabidops
4	209.2	2.2	480	10	AB001426	AB001426 Mus muscu
5	206.6	2.1	581	6	AX321334	AX321334 Sequence
6	177.8	1.8	7524	3	AF220353	AF220353 Drosophi1
7	176.4	1.7	3692	8	AB062739	AB062739 Arabidops
8	167.6	1.7	3273	8	AB062738	AB062738 Arabidops
9	162.6	1.7	3599	8	AY039966	AY039966 Arabidops
10	161.2	1.7	1723	8	AF372930	AF372930 Arabidops
11	157.2	1.6	3734	10	MUSKIF3BA	D26077 Mouse mRNA
12	154.4	1.6	69523	2	AC094844	AC094844 Rattus no
13	153.2	1.5	2213	3	SUSPKIN2A	L16993 Strongyloce
14	149	1.5	2759	9	AB072777	AB072777 Meaca fa
15	148	1.5	3761	5	XLA9839	AT09839 Xenopus 1
16	147.8	1.5	2109	9	AF041853	AF041853 Homo sapl
17	147.8	1.5	2565	8	AF272749	AF272749 Zea mays
18	146.2	1.5	2282	10	MUSKIF3	D12645 Mouse mRNA
19	144.8	1.5	2500	3	AF149286	AF149286 Caenorhab
20	142	1.5	2548	3	TTTH244020	AJ244020 Tetrahymc
21	141.8	1.5	3688	9	HSKICMR	X65873 H.sapiens m
22	140.2	1.5	3341	9	BC009353	BC009353 Homo sapl
23	138.2	1.4	501	6	AX321150	AX321150 Sequence
24	135.2	1.4	2888	8	AF272759	AF272759 Zea mays
25	135.2	1.4	3705	9	HSAA271784	AJ271784 Homo sapl
26	135.2	1.4	3771	8	AF179308	AF179308 Homo sapl
27	134.4	1.4	2598	8	AF272750	AF272750 Zea mays
28	133.6	1.4	4348	9	AF071592	AF071592 Homo sapl
29	133.4	1.4	2100	5	XLA311602	AJ311602 Xenopus 1
30	132	1.4	4724	9	AB002357	AB002357 Human mRN
31	132	1.4	145414	9	HSAA392M18	AL121897 Human DNA
32	131.2	1.4	18780	2	AC019734	AC019734 Drosophi1
33	131.2	1.4	94440	3	AC006073	AC006073 Drosophi1
34	131.2	1.4	174652	3	AC092241	AC092241 Drosophi1
35	131.2	1.4	261383	3	AE003631	AE003631 Drosophi1
36	130.8	1.4	3791	10	MM086090	U86090 Mus musculu
37	130.4	1.4	136366	2	AC008410	AC008410 Homo sapl
38	130.4	1.4	157375	2	AC025747	AC025747 Homo sapl
39	128.8	1.3	903	33	AC076014	AC076014 Glardia 1
40	127.6	1.3	8743	9	AF279865	AF279865 Homo sapl
41	127.4	1.3	3432	10	AB008867	AB008867 Mus muscu
42	127.2	1.3	2229	3	U00996	U00996 Strongyloce
43	127.2	1.3	4125	9	AF241316	AF241316 Homo sapl
44	126.6	1.3	3675	3	SODKINEA	J05258 Squid kines
45	123.6	1.3	468	3	AF015713	AF015713 Dictyoste

ALIGNMENTS

RESULT	1					
AF027728	AF027728	9610 bp	MRNA	linear	VRT 15-NOV-1997	
LOCUS	Xenopus laevis kinesin-related protein (XCENP-E)	complete				
DEFINITION	cds					
ACCESSION	AF027728					
VERSION	AF027728.1	GI:2586070				
KEYWORDS						
SOURCE						
ORGANISM						
	African clawed frog.					
	Xenopus laevis					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
	Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;					
	Xenopodinae; Xenopus.					
REFERENCE	1 (bases 1 to 9610)					
AUTHORS	Wood,K.W., Sakowicz,R., Goldstein,L.S. and Cleveland,D.W.					
TITLE	CENP-E is a plus end-directed kinetochore motor required for					
JOURNAL	metaphase chromosome alignment					
MEDLINE	Cell 91 (3), 357-366 (1997)					
REFERENCE	2 (bases 1 to 9610)					
AUTHORS	Wood,K.W. and Cleveland,D.W.					
TITLE	Direct Submission					
JOURNAL	Submitted (01-OCT-1997) Cell Biology, Ludwig Institute for Cancer					

FEATURES
Research, 9500 Gilman Dr., La Jolla, CA 92093-0660, USA
Location/Qualifiers

gene
1..9610
/organism="Xenopus laevis"
/db_xref="taxon:8355"

CDS
135..8999
/gene="XCENP-E"
/note="kinesin-related protein"

/product="kinesin-related protein"
/protein_id="AAC60300.1"
/db_xref="GI:2586071"

1..9610
/organism="Xenopus laevis"
/db_xref="taxon:8355"
135..8999
/gene="XCENP-E"
/note="kinesin-related protein"
/product="kinesin-related protein"
/protein_id="AAC60300.1"
/db_xref="GI:2586071"
/translation="MSEGDAAVCVRVPLIQROGQANLQMKANNTISQVDTGKS
FNDRVFNSHSTSOIYOEAIVPIIRALQSTNITFAGQTSSTGTYMMGPNISG
IIPAOIEVFKIOEIPNRELLPSTMEIYNETVDTLDCDRKKPRLIEDFPNRY
YVADLLEIVYVEHVIQWIKKGNHGYEIMNDHSSRSHITFPMIYESDRDPT
NSCNDGAVVSHINVDYLAERASQTEGVRKMGKNSINSLFLLGOVYKISDQ
QAGCFINRDSKLRITLQNSLQGNKAVIICITTPYSPFETISTLOFASTAHVNT
HNEYIDDEALKRKKEIIDLKQLENNESSEKKAQMADEHQTALAEIKHNP
REDRIWH/TVIVASQESQDQVRKRVWAPKIQNSIAGSVDPMLSLPLG
NSKKARSDMPSPPEIDSVCTEFSDPDALSMNSGIDEMNLASKVTHREKSL
HOSMIDGQISDSVQFHDSSKRNQOLYLPKDGDAECNKAQFEKETSILQOQOSE
LQSDGYNASNSDLODSVDGKRLSSHDCTIHRMMLQKIVDLEEFLENKSEND
KQNSQDPMESIOICEAIIAEKANAELIARMPNIIENETLREIADLERST
QEOIARERKADNLOKRVNINOKEKASLEKEIEIYFOLKQSLASQVTRDQKE
CSFRLSENLEIKEMEDISNMYNOKEKASLEKEIEIYFOLKQSLASQVTRDQKE
NEINYLNGIAGVPRDLISVLEKKESEFSEKQLEKALENBYCLOSEYFEL
PNECTKQISASBEIIMLKQEGEHASITISKOETIMQOSQIQLDVEYTHQS
KNOETQOELERKMHNDLFEKIRNKSEADILREMNKGTMESEVYIATKHEL
EETIRKQOELERKMHNDLFEKIRNKSEADILREMNKGTMESEVYIATKHEL
LATERNNINWCLTERNSLBOVDIANTOLOSIOAQSTESDLQKQKODIEEYVIL
LEBELKHLIDLSQSLIKLOLENIETEKIQTIOEMKNIITERNELDIPEDLKAE
HOSLQKQSENIEOSIETODELRAQOELBOQLOVSEKNIITERNELDIPEDLKAE
ANQEVSLGEVNSLOSEMLRGEEDLOSCAKVASELELIRAHVSEENLEIYKIL
NGLEKEIIGKSESESVLKSMLEKEDNNKLKQAEYSSEKQNOFSELPVSGSKVL
DEIEVLAQKAAERLEIKRODIFELVOTANNIYVKGLEPIPLADHEDSDIRSE
EMILKVEKTERNOYLLERQOEKELLSKLEIILKEMETVILLKDILOQLESLS
PLEKILLILBELHQKTEQEKLEHEKNELEQAVELKGEVSHLMKSIKESLSL
OHEKHDQOQLLQOMQVYQEKKELOOTHEHILAEVDHLKENIELGFKPNDAQ
KITTEQCLNENKLEQSOHROCEIEELMKSIKQESALELTKELQALQOQKVOEL
VWLEMEELKNSORTVIAERDLODLRESVMSIETQDDLKQALQOQKVOEL
TSQISVLOEKISLENOMLYNAVATYETISEDDQNOKQHFSEIETLSIKERKE
ALBOEKQKADARKTIDITEKISNIEQOLQOATNLEKTYERRESILOKQALAT
EHLRETLKSDIALGKEDRODAANKYALTEKSSLEQOINENVTTLKGBGEKET
FYLOPSSKOOSQOMELRESLTKDLOLEAEKISATNETIKUTAKISLSEELT
ONASIIENAVSEKENLRHSKOOLVSELQSLTLKSRPRAQSKREKIDAVKIASL
AERIKLITKEMDFPDSKESLOESSHSEELCYKTELOMLKQKREKIDAVKIASL
EDBELQHLISLKEQDLOJOMELRNEKIRANYLEKQKIDNEKELISYRLQNOQEE
DVAERMDILSEKNOEIOELMEKISAVYSEBHTLSEJLSEKELREAHKHNKE
SLSSTLSRSGISLOTEHYKNTLOLOTLINKKVVYRYAAVEKSHLSEKIDKLAQ
KRHDELRLQLOLLEOHGRKMSDSAEELCEIELELLEFKANNIIOGVQDESEVO
VILNVOGSIQOELEHKKQPMOLEEFGVLAHAKKISEGQOENRISTQOILKRI
LKVAVOSKIOREITVYLLNQFPAKLOEKQNEKLEMRMEHGSASVMEENALGLI
ITKQDNQIAAKKEITENLQKVAAGAVPYKEEDNTKTYVATIMEKIYSATQOET
AYKISCLDEKEGRLKELRAQANDNTTVNPKKRAKASDHASHTSGSSHSGSP
AMLVYQSKALERELESHYKKYHHLSRIMSSSEDDKRAKASDHASHTSGSSHSGSP
HKTETGRGVALTERSEMPSLHSGPKKSSSRKRVNSNREISITQVMSPPCKQKH
VOENLSPKGLKRAKSPRSMPTQHTISPKGLHKNLLESTLFPNLSPPCKQKH
KETAPCKKTS"
1929..1955
/gene="XCENP-E"
/note="nucleotide sequence is absent in another XCENP-E
cDNA"

BASE COUNT 3536 a 1688 c 2100 g 2286 t
ORIGIN

Query Match	99.8%	Score 9610	DB 5	Length 9610
Best Local Similarity	100.0%	Pred. No. 0		
Matches 9610	Conservative 0	Mismatches 0	Indels 0	Gaps 0
9	gagtcgataagctagctcgagcgaaggaatccaactggtatagaagaactgaac	68		
1	GAGTCGATAGGCTAGTCGCGGAGGCAAAATCAACTGTTATAGAAACCTTGAACC	60		
69	gccgcgaagaaggaactaagtgacagagagagagagcggctgcggtaccgaattccac	128		
61	GCCGCGCAAAAGGACTTAAGTGACAGACAGAGAGGAGCGGTGCGGTACGATTCTCCAC	120		
129	taatcgctcaaatatccgagagagagatgcagttaaagtgtgtgagggcttgccgc	188		
121	TAATCGGCTCAAAATGTCGAGGAGATGCAAGTAAAGTGTGTGAGCGCTTGCCGC	180		
189	ttatacgaagaagaagaaggagatcaagccaactgcattgaagaagctggaacaaccca	248		
181	TTATACGAGAGAAACAAAGGAGATCAAGCCAACTGCATGSAAGCGTGAACACACA	240		
249	ttcccaagttgatggagacaagttctcaattccgactggtatattaaattccacgaat	308		
241	TTTCCCAAGTTGATGGACAAAGTCTTCAATTCGATCGTATTAATTCACGAT	300		
309	caacaagtaaatattcaagaagaatagcagtaaccatcatagatagcttgcaaggat	368		
301	CAACAAGTCAAAATTTACCAAAATACAGTACTATCATGATGATGATGATGATGATGAT	360		
369	ataatgcacaaattttgacatagcagacacatcttcaggaagacgtacacaaatgaag	428		
361	ATAATGCACAATTTTGCATATGAGACAGACATCTTCAGGCAAGACGTACAAATGATG	420		
429	gaacaccaaattcttgagcgaataatcccaagccaataagaagaattttaaattttc	488		
421	GAAACCAAAATTCATTGGGATATATATATATATATATATATATATATATATATATAT	480		
489	aggagataccgagacagagagattctcttaagaagttcttatatagagattacaatga	548		
481	AGGAGATACCGAACAAGAGATTCTTAAAGTTCTTATATATATATATATATATATATAT	540		
549	ctgtgaagaacctacgtgtgtgtagcagaagaagaagcccttgaattccgagagatt	608		
541	CTGTGAAACACCTACGTGTGATGACAGAAAGAAAGCCCTTGAAATTCGAGAGATT	600		
609	taataagaagctgtatgttgcctgaactgtagaagactgtgaatggttcttgaacatg	668		
601	TAAATGAAACGCTGTATGTTGCTGACTGAGTAAGACTTGTATATGATGCTGAAATG	660		
669	taatacagtgatcaaaaagggtgaaaaaaacagacattatggagagacataaagaatg	728		
661	TAAATCACTGATCAAAAAGGCTGAAAAAAGACATTAATGAGACACTAAATGATG	720		
729	atcatagaagtcgttccatatacaataattagaatgtgttgaagaagccgaagaagaatg	788		
721	ATCATAGTACGTCGTCACATACATATTTAGAAATGTTGTTGAAGCCGAGACGAATG	780		
789	atcccaaatatcaagaagactgtagatgagctgtcatgtatctccacttgaattgttag	848		
781	ATCCCAAAATTCAGAGACTGTGATGAGCTGTATGATGATGATGATGATGATGATGATG	840		
849	atcttcctgagctgaagaagaagcaaaccttgaagctggaaggtgtgaagactgaagaag	908		
841	ATCTTCTGCGAGTGAAGAGCAACCAACTGAGAGCTGAGAGCTGAGAGCTGAGAGAG	900		
909	gctgaagaatcaaccgagctgttataccttggacagcttataagaagcttgaagag	968		
901	GCTGGAACATCAACCGAGCTTGTATCTTGTGACAGGTTATTAAGAGCTTACGAGC	960		
969	gccagagctggtggttataactacagagacagcaaatcccaagaatttcccaaat	1028		
961	GCCAGGCTGTGATTTATTAACACAGAGACAGCAAACTACCAAGATTTCCCAAAAT	1020		

QY 1029 cattggaggaaatgctaaacggtataatcttcacaaatlaagccagttcttttgatg 1088
DB 1021 CATTTGGAGAAATGCTAAACGGTTATATATTTGACAAATTTACGCCAGTTCTTTTGATG 1080
QY 1089 agacttaagtacactcagcttgcagctacgtccaaactgtggaataactcccatg 1148
DB 1081 AGACTTAAGTACACTTTCAGTTTGGCAGTACCTCCAAACGTGTGAGAAATACCTCCCATG 1140
QY 1149 ttaatgaggtccctbga tga tgaagcgtctctaaaaagttacagaaggaatcttgat 1208
DB 1141 TTAATGAGTCCCTGATGATGAAAGCGTTGCTAAAGGTACAGAAAGAAATCTTGAT 1200
QY 1209 taaagaacaatlaaggaatttagatcactcgtctgaacaagaagctcaagcaatggtc 1268
DB 1201 TAAAGAAACAAATTAAGAAATTAAGATCATCGTCAAAACAAAAGCTCAAGCAATGCT 1260
QY 1269 aagaagagcatcacagcttgcagctgaatcaacaactacacaagaagagagaatga 1328
DB 1261 AAGAAGAGCATACACAGTTGCTAGCTGAATCAACAACTACCAANAGAGAGAAAGATA 1320
QY 1329 gaatagtcacttgacaatatctgtctgtctcatalcccaagaatctcaacagagacaa 1388
DB 1321 GAATATGCGACTTGACAAATATTTGTTGCTTCATCCCAAGATCTCAACAGACCAAA 1380
QY 1389 gggtcacaagaaacgaagaatcagctgggcgacagaagaatcccaaatatgattacatg 1448
DB 1381 GGGTCAAAAGAAACCAAGATGATGCGGCGCAGAGAAATCCAAATATGTTTACATG 1440
QY 1449 ctctctggtcttctgacttgatgactatccagattactgcaatttagaagaagaag 1508
DB 1441 CTCTCTGTTCTTGCTTGATGATGCTTCCGATTTACCTGGCAATTTTGAAGAGAG 1500
QY 1509 caaagctcctgcacagctctcatalcccaagaatltgaactcgtctgtacagaagttc 1568
DB 1501 CAAAGTCTCTGCATGCGCTTCATTTCCAGAAATGATGATGCTGTTTGTACAGAGTTT 1560
QY 1569 ctgattcttgaagccctctccca tga tgaacagcaatggaatagatgcagaatgtaact 1628
DB 1561 CTGATTTTATGAGCGCCCTCCATGATGAGCAATGAAATGATGCGAATGGAAT 1620
QY 1629 tagccagtaagtaactcacagaagaagaacatcacttcaactaataatgataagcttg 1688
DB 1621 TAGCCGTAAGTAAGTACACAGAGAAAGACATCACTTCAATCAATGATGACTTTG 1680
QY 1689 gacagaattctgcagctgtctcagttcactgattctcttaagaagaacccagctcaat 1748
DB 1681 GACAGATTTCTGACAGTGTTCAGTTTCATGATTTCTTAAGAAACCAAGCTACAAATAC 1740
QY 1749 tccccaagaactcctggtgata tggctggaatgcagaagaagctctctttgaaagaatga 1808
DB 1741 TCCCCAAACACTCTGCTGATGATGGCTGAATGCGAAAGGCTCTTTTGAAAAAGACTCA 1800
QY 1809 caagcctcagcaacaacacagctcaagaagagaagaagaagaagaactgttacaagct 1868
DB 1801 CAAGCCTCCAGCAACACTACAGTCAAAAGAGAGAAAGAAAGAACTTTGTCAAGCT 1860
QY 1869 tccagctcacaagtagcagaacttggaaagcagctttagtctcaaaagctaaaaactcaga 1928
DB 1861 TCGAGCTCAAGATAGCAGAACTGGAAGACAGCAGTGTAGTCAAAAGCTAAAGAACTGAG 1920
QY 1929 tgggttaacaactcagagaagcattccatcaatgctgaagtccaacagatgttgaagaag 1988
DB 1921 TGGTTTCAAACTCGAAGAGCATTTCCATTAATGCTGAAGTCCAAAGATGATGTTGAAGG 1980
QY 1989 aagctgtgagaagaagaatgctagctctgtgagaactcgtgtacaatgctcaacaagct 2048
DB 1981 AAGTTTGTGAAGAAAGAAAGTCAAGTCTTGAGACTCTGTTTCAATGCTCAAGAGTG 2040
QY 2049 acctacaagatagttctgttga tgc tgaacgltcaagcagctcccatgattgattag 2108
DB 2041 ACCTACAGATAGTTCTGTGATGATGTAAGCTAAGCAAGCTCCCAATGATGATGATAG 2100

QY 2109 aaccacagaataatgctcgaaacaaaga tgc tga ttttagaagaatttactgaaaacct 2168
DB 2101 AACACAGAAAATGCTGGAACAAAAGATCGTTGATTTAGAGGTTTATTGAAACCTTA 2160
QY 2169 acaagaagaag tgaagatgataaacaanaaactcttcgacagaatctttagaagatg 2228
DB 2161 ACAAGAAAGTGAAGATGATGAACAAAATCTTCTGACGAAGATTTATGAGAGATATTC 2220
QY 2229 agctatgtgaagctata tggcagaagaaggcaaatgcacttggaggaaactgacatg 2288
DB 2221 AGCTATGTGAAGCTATTAATGCGAGAAAGCAATGCACTGAGGAACCTGCACTTAATG 2280
QY 2289 ggaataatttgacaatattatcttagaagaatgaactcraaaagggaatctgagaatc 2348
DB 2281 GAGATTAATTTTGAACAAATTAATTTCTAGAGATGAACCTTAANAAGGAATGCAAGATC 2340
QY 2349 tggaaagctcacacaaaggaatcaagaacaaactgagttgaaatcttgagaagaagaa 2408
DB 2341 TGGACGTTCACTCAAGCAATCAAGAAACCAATGAGTTGAAATTTCTGAGAGAGAA 2400
QY 2409 ctcaaaaaagaacagaggcaaacata tcca tgaagat tggcaglttaagaagaat tga 2468
DB 2401 CTCAAAAAGAAACAGAGGCAACCAATCCATGATGAGATTGGCAGTTTAAAGAAATTAATG 2460
QY 2469 aaaaatgcagaatgataacataacatacttgaagaaga tccagaacataaacaacatc 2528
DB 2461 AAAATGAGAGATGTACAACTAAATCTTGAAGAGATCTAAGAACTTAANAACAAATTC 2520
QY 2529 tgaagaagcagaagaatcacttgcagaatctaaagaagaacgcagataactctgagaaga 2588
DB 2521 TGAAGAGCAGGAATTCACACTTGCAGAAATTAAGAAACCCGCAAGTAATCTGCAGAAA 2580
QY 2589 aagtaacgaatcttgaatctcgttctca tgggtga tag tgaagaacactgttgaaga 2648
DB 2581 AAGTACGAATTTTGAATCTCTCGTTTCCATGGGTGATGAGAACTCTGGAAGAA 2640
QY 2649 tcttcaactgaagaacatctctcttctgtatgctgaagctgagctcgaagaag 2708
DB 2641 TCTTTCAACGAAGCAATCTCTTCTGATGCTGAACCTGACCTCGGAGTCTCAGAGG 2700
QY 2709 aatgtctctcccaagaag tgaagaatc laagct tgaagagaagaat tgaagaacatca 2768
DB 2701 AATGTTCTTCTCCTCAGAGAGGAATTTAGAGCTGAAGAGAAATGAGACACATCA 2760
QY 2769 actgtatacaatcaaaaaagaaggctgctctctgttcttgaagaagcagctggaactga 2828
DB 2761 ACTGTTCAATCAAAAAAGAAAGGCTCGTCTTGTGTTGAGAAAGCTGGAACCTGAA 2820
QY 2829 aatcaaacatacaagaagaatggaagctgatttgaagaagaagttgcaagaagctttatg 2888
DB 2821 AATCAAACTGCAAGAAATGGAAGCTGATTTGCAAGAAAGGTTGCAAAAGTCTTTAATG 2880
QY 2889 agatlaactactaaatgagctctctgcaagaagaagctcccaagaagatctctctctg 2948
DB 2881 AGATTAACCTTAATAATGAGCTTCTGCGAGAAAGAGTCCCAAGAGATTTCTCTCTG 2940
QY 2949 tgaatbagaagaagaagttctcgaagltctcaagaagcagcttgaagaagca t tgaaga 3008
DB 2941 TTGAATTAAGAAAGAGTTTCTGAGTTCTCAAAAGCAGCTTGGAAGAACTTGGAAGAA 3000
QY 3009 aaaaatgcttggagaatggaatgagcttgcataccaagaatacaatcttgcgaagaag 3068
DB 3001 AAAATGCTTGGAGAAATGAAGTACCTTATTCAGAAATCAAAATTTTTCCAAAATGAG 3060
QY 3069 tgaatgcttgaagaataatcaatcagcaagctctcgaagaatlaatgcttcaagaag 3128
DB 3061 TTGAATGCTTGAAGAAATCAATCAGCAAGCTTCTGAAGAGATTAATTTTAAGCAAG 3120
QY 3129 aaggagaaca tctcgaatctataagaacaagaagatlaatcaltgacagaagaatctg 3188
DB 3121 AAGGAGAACTTCTGCTATTAATTAAGCAAAAGAGATTAATCAAGACAGCAATCTG 3180
QY 3189 agcagaatttacaactgacagcaggaag tgaacacacacagctcaaaagtgcaagcagctg 3248

```
|||||
Db 3181 AGCACAATTTTAACTGACTGACGAGTACACACACACACAGTCAAAATGACAGCAGACTG 3240
OY aagagcaactacttgagatgaaagaaatgcatgactcttltgaaaaatatacagaa 3308
Db 3241 AAGACCAATFACCTGGAGATGAGAAATGCATGATGATCTTTTGAAGATATACAGAA 3300
OY 3309 acaaaagtgaagctgaagacctttaaagaaatggagaaccttaaggcactagagat 3368
Db 3301 ACAAAAAGTGAACCTGAAGACCTTTTAAAGAAATGAGAACCTTAAAGGCATATGAGAT 3360
OY 3369 cgttggaatgaagattgtgcacacaaacatgaaacttgaagaactaaagataag 3428
Db 3361 CGTGGAAATTAAGATTGCTGACACAAAACATGAACTTGAAACAACTTAAGGATTAAGG 3420
OY 3429 agcaactgtctcatgagaaaaatacttlttcaagcaatgacagatatlattccgatta 3488
Db 3421 AGCAACTGCTCATGAGAAAAATFACCTTTTCAAGCAATGCAGCTATATTTCCGATTA 3480
OY 3489 caactcttcagactcgctctccctcaaaatagttgaaggagactcaagaccaca 3548
Db 3481 CACCTCTTTCAGACTGCTGCTCCTCAAAATTAAGTTGAAGGGAAGCTCTCAAGACCCA 3540
OY 3549 tagaaatcaatgactacccaaatlaatatagcccttgctacagaaagaaacaaatlatg 3608
Db 3541 TAGAAATCAATGACTACACCAATTTAATAGCCCTTGCTACAGAAAGGAACAAACATTAAG 3600
OY 3609 tgtgtctagagactgaaagaaacagctctcaagagcaagttattgatttgaacactaac 3668
Db 3601 TGTGCTAAGAGACTGAGAAACAGTCTCAGAGAGCAAGTTATTTGAACACCTCAAC 3660
OY 3669 ttcaagttctcaagacaaagatgaaagatcgtatccagagagccaaagcaagact 3728
Db 3661 TTCAAAAGTCTTCAGACCAAAAGTATTGAAAAGTCTATCTCCAGAGCCAAAGCAAGACT 3720
OY 3729 tggaaagaaagagagttaaactgtcttggagatggaaactgaactgaaggacaactatg 3788
Db 3721 TGGAAAGAGAGAGGTAAACCTCTTTGGAGATGGAACACTGAAGGACACCTAACTG 3780
OY 3789 aactcaagctgtctatagagaaatgcaagctggaagaaactggaagcttaagaanaactgc 3848
Db 3781 ACTCAAGAGCTGTATAGAAATTCAGAGTGAAGATCTGGAAGTTACAGAAAACCTCC 3840
OY 3849 aaacacttcaagaaagatgaaataattactatagaaagaaatgagcttcagaccact 3908
Db 3841 AAACACTTCAAGAGAGATTAATAATTTACTATAGAAAGATGAGCTTCAACCAACT 3900
OY 3909 tcgaagacactgaaagcagagcatgataagcctaaacaagaaccttaagtaaacattgagc 3968
Db 3901 TCGAAGACCTGAAAGCAGACATGATAGCTTAACAAACAGACCTTAAAGAAACATTGAGC 3960
OY 3969 agtcaaatgaaacaaagatgaatgaatgaagctgcccagagaaagctaaagaaacagaagc 4028
Db 3961 AGTCAAAATGAAACAAAGATGAATTAAGGGCTCCCAAGAGCTTAAGGAACAGAGAGC 4020
OY 4029 aactggttgaacttagacaacagactttagatgtctgtlaggaattcatcccaa 4088
Db 4021 AACTGTTGATFACCTTAAACAAACAGCTTTAGATTGTTCTGTAGGAATTTCAACCAA 4080
OY 4089 accatgataagctgtcaaccagaaagatgtcaatgggtggatgaatgaatgaatgaatgaat 4148
Db 4081 ACCATGATGCAAGTGTACACAGGAAAGGTGTCATGGGTGAAGTTAAATTCCTTCAAA 4140
OY 4149 gtgaaatgctgctgtggaagagatgagcttcaaacatcttgtaagcatgtttcag 4208
Db 4141 GTGAAATGCTGCGTGGTGAAGAGATGAGCTTCAACATCTTTGTAAGCACTTTAGTTTCAAG 4200
OY 4209 aactgagacactcgtgctcatgtaaaactgtgagggagaaacacttgaataacaa 4268
Db 4201 AACTGAGGACTACTGCTGCTCATGTAAATCTGTGAGGAGGAAACCTTGAATCACA 4260
OY 4269 aaaaactcaatgacctgaaagagatatitggcacaactgaaagaagcgaagtgtga 4328
|||||

Db 4261 AAAAACAATGAGCTTGAAGAGAGATATTTGGCAAACTGAAGAAAGCAAGTGTGA 4320
OY 4329 agtccatgttggagaaatctaagaaagacaacaaataagctcaagaagcagagagat 4388
Db 4321 AGTCCATGTTGGAAATCTTAAGAGACACACATTAAGCTCAAGAAACAGAGAGAAAT 4380
OY 4389 attcagtaaaagaaatcaatcactcagtttagaagaggtgttcagtggttcaagaagcttg 4448
Db 4381 ATTCTAGTAAAGAAATCAATATTCAGTTTACAGAGGTGTTCAAGTGTTCACAGAACTGG 4440
OY 4449 tagatgatatagagctctgaaagcacagctaaagagaaagaaagcctgaaataa 4508
Db 4441 TAGATGAGATGAGCTCCTGAAGCACAGCTTAAAGGAGAGAGAAAGCTGCAAAATTA 4500
OY 4509 aggatagagatacttgaacttgaacttgaacaaactgcaaacacaaatltgattgaaggaaat 4568
Db 4501 AGGATAGGATTTATTTGAACCTTGACAACTGCAAAACCAATTTAGTTGAGGAAAT 4560
OY 4569 tggaaactcatgcaagctgacaaatgagaaagacagcatgtgacggttctgaaagaa 4628
Db 4561 TGGAAACTCCATTGCAAGCTGACATGAGAGAGACAGATGATGAGCGTTCGAAGAAA 4620
OY 4629 tggagataaagctcttgagaaagaaacttgagcgaatcaglatcttaactgaaagattgc 4688
Db 4621 TGGAGATTAAGGTTCTTGAGAAAGAACTTGAGCCGAATCATATTTACTGGAAGATTGC 4680
OY 4689 aagaagaaagctggaactgtctaaacaaacttgaatccctcagaagaaatgagagact 4748
Db 4681 AAGAAGAAAGCTGGAACCTGTAAACAACTTGAATCTCTCAAGAAAGATGAGAGAGCT 4740
OY 4749 ccgttctatlaaagaatgaactgtcaacagaaagcttagaaagcttgcagtgaaagactta 4808
Db 4741 CCGTCTATTTAAAGATACCTGCAACAGAAAGCTTGAAGAGCTTGCTGTAAGAAACATTA 4800
OY 4809 ttctaaagagagataitgaacaaacccaaagcatcattcagataactcaagctcagctgc 4868
Db 4801 TTCTTAAGAGATTTTATGACACACACCTTAAGCATCATTTAGATATCTCAACCTCAGCTGC 4860
OY 4869 agaaacacagaaagagctacagttgctgaagatcttgcgaatcgctgccttcgaact 4928
Db 4861 AAAAAACAGCAAGAGCTACAGTTGGCTTAAGAACTTGCATGCTGCTCTGCAATTT 4920
OY 4929 gtccaataactaaagaaagaaacctcttgcagatttgcattccctctgaaagaaaga 4988
Db 4921 GTCCATAACTCAAAAAAGAAAGAACTCTGCAAGATTGCTGATCTCTTGGAGAAAAAGA 4980
OY 4989 tatgttataactgaagaatgcaatcaaaaaactaaatgaacagaaagaaatctacatg 5048
Db 4981 TATTGTTATTAATCTGAAGATTTGCATCAAAAAACTTAAGAACAGGAAAAATTACTACATG 5040
OY 5049 aaaagatgaacttgaaacagctcaggttgaagctaaagtgagagtggaacatctgata 5108
Db 5041 AAAAGATGAACCTGAACAAAGCTCAGGTTGAGACTTAAAGTGCAGAGTGAACATCTGATGA 5100
OY 5109 aagatgataagaaatcgaaagctcacttgaagcttccttgcagaaatgaaagaaatgatactg 5168
Db 5101 AAGATGATGCAATGCAAGAGCTCTCACTTGATGCTCTTGCAGATGAGAAACATGATCTG 5160
OY 5169 aacaaacttctgctctttaaacagcagatgcaatgaatgaatgaatgaatgaatgaatgaat 5228
Db 5161 AACAACTTCTTGCTGTTAAACAGCAAGATGCAAGTATGATCTCAAGAAAGAAAGAGC 5220
OY 5229 tgcacaaaccaaatagaacacttaacagctgaggtgagacatctaaagaagatattgagat 5288
Db 5221 TGCACAAACCAATGAAACCTTAACAGCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGT 5280
OY 5289 tgggttgaattttaaaaaatgaaagcgaagcaaaagacaaactaaagaagcatgtctgtaa 5348
Db 5281 TGGGTTTGAATTTTAAATTAAGAGCCGACCAAAAGCCACTTAAGAGCAATGTCTGTAA 5340
OY 5349 atgagataaagaaacttgcagaaagcagacagacttcaatgtgagatagaaagctta 5408
Db 5341 ATGAGATTAAGAACTTGAAGCAAGCCACAGACTTCAATGTGATGATGAAGAGCTTA 5400
|||||
```


QY	5409	tgaagagcttaagataaagagtcagcgcgtgtaaaccttaagaatctgacaaaag	5468
Db	5401	TGAAAAGGCTTAAAGGATTAAGAGTCAGCGCGGTGGAAACTTTAAAAAACCTGAAACAAAAG	5460
QY	5469	ttaataacttbaaccaagaagatggaaatgtgtatgtctgtgaaaatggaggaattgaaaata	5528
Db	5461	TAAATTAATCTGAAACCAAGAAATGGAAATGTTATGCTGGAAATGGAGCAATTTGAANAATA	5520
QY	5529	gccagagagactgttaattgtctgaaagaggaccagctgtcgaacgacacttaaggagagtgctg	5588
Db	5521	GCCAGAGAGACTGTAAATGTCGTGAGAGGAGACAGCTGCAAGACACCTTAAGGGAGAGTGTGG	5580
QY	5589	aaatgtccattgaaacctcaagaatgatacttaagaagcttaagaacatttgcagacgcga	5648
Db	5581	AAATGTCCATTGAAACCTCAAGATGATCTTAAGAAAGGCTTAAAGACATTGCGAGCGAGA	5640
QY	5649	aagataaagtctcagaagactgaacctccagatctctctccctgcaagaaagatctctctc	5708
Db	5641	AAGATTAAGTTCAAGAGACTGACCTCCAGATTTCTGTCTCTGAGGAAAAAGATCTCTCTTC	5700
QY	5709	tggaaaatcagatgcttataatgtgtgcaactgtgaaagaaaccttaagcgaagagatg	5768
Db	5701	TGGAAAATCAGATGCTTTATTAATGTTGCAACCTGTGAAGAAACCTTAAGCGGAAGAGATG	5760
QY	5769	accggaacacagcttaagcaaacacctgtcttcacgaanaatgaaaccttgatcgtctctaa	5828
Db	5761	ACCTGAAACAGCTTAAGCAACACCTGTCTTCAGAAATTAACACTTTTGTCTCTTTTA	5820
QY	5829	aagaaaggaattgtcatctgaaacaagcagagaaaggacaaagctgatactgtccagaaaa	5888
Db	5821	AAGAAAAGGAATTTGGCTTTGTGAAACAAGCAGAGAGACAAACCTGATCTGTCCAGAGAAA	5880
QY	5889	caataagatacacaagagaagatacgaatataagaagaacagttactttaaagaacacaca	5948
Db	5881	CAATAGATATCACAGGAATAATATCAAAATATGGAAACAGTACTTCAACAAGCACACA	5940
QY	5949	atttaagagaaaccttctacgaagagagagagttctaccagtgtaaggagacagtgctt	6008
Db	5941	ATTTAAGGAACACTTTTATGAAGAAGAGAGTCTTATCCAGTGAAGGACMACTGGCTT	6000
QY	6009	tgaacacagaacaccttaaggaaaacttgaagagcaagaagacttgatctgtgttaaatg	6068
Db	6001	TGAACACAGAACACTTTAAGGGAACATTTGAAGAGCAAAACACTTTGGCATTTGGTAAATAGG	6060
QY	6069	agcagagagagatgtaagctgcgcacataaagtaagctcttcacgaagaagatgtctctc	6128
Db	6061	AGCAGAGAGAGAGATGAAGCTGCCAATAAAGTATAGCTTTACAGAAAAAGATGCTTCTTC	6120
QY	6129	tagaagaacagatcgaatgaaatgttactactcttaagaagvgtagvgtgaaaaagaa	6188
Db	6121	TAGAAGAACAGATCAATGAATAATGTTCATACCTCTAAAGANAGTGAAGGTGAAAAAGAGA	6180
QY	6189	gcttctacaccttgagaagaccttgaagcagacgacgtctctcccaatggaagaagctcagag	6248
Db	6181	CGTTCTTACCTTCAGAGACCTTCAACAGCAGCAGTCTTTCTCCAGATGGAAGAGCTCAGAG	6240
QY	6249	agctcttaagaccaagaatlttgcaatctgaaagagcgcgagaagaagataaagtgaagcta	6308
Db	6241	AGCTTTTAAAGACCAAAAGATTTGCAGTTGGAAAGGCCAGAAAGAGATTAAGTGAAGCTA	6300
QY	6309	ctaatgaaataaagaatctctacatgcgtcaagaatctctctctcttagaagaagaatcttctaga	6368
Db	6301	CTAATGAATAAAGAAATCTACAGTCGTAAATATCTCTTCTCTTAAGAGAGAGATTTCTTCAGA	6360
QY	6369	atgctagacatcttgaatgtaagctgttaagcgaagaaggaaaaccttcgcacatctgaagcagc	6428
Db	6361	ATGCTAGCATTTTGAATGAAGAGCTGTAAAGGAAGAAAGAAACCTTGCCATTGCAAGACAGC	6420
QY	6429	aactgtgttcagaatctvgagcagctatacagctgacatlaaagaatagagacacatgcatltg	6488
Db	6421	AACGTGTTTACAGATTTGAGACACCTTCTACATGACATTAAGAGTGAAGACCATGATTTGG	6480

QY	6489	cgcaatcttcaaacgcgzaaagga tgaagc tgttaa taaatbaagccag tccgcgttgaana	6548
Db	6481	cccaatcttAAACGGGAAAGGATBAACCTCTTAATATAAATAGCCACTCTCGCTGAAGAA	6540
QY	6549	taaaatccctcaaaaagaga tga tgaatcttcaagat tccaaggaatcccttgcagaagc	6608
Db	6541	TTAAATCTCTGCAAAGAGAGATGATTAATTCAGACGTTCCAAAGGAAATCCTTGCAGAAAC	6600
QY	6609	agcttcccatctcaag tgaagat tga taca taaagac tgaacttcaat tgc tcaagc	6668
Db	6601	AGCTTCCCACTTAAGTGAAAGTATATGTCATATTAAGCTGAACTTCAAAATGCTCAAGC	6660
QY	6669	aacgaagaagaacatcaacaacaacttgcagaagaagttaa gagaag tga tgaactat	6728
Db	6661	AACAGAAAGAGACATCACACAACTTTCAGAGAAAGTTAAGAAAGATGATGACTAT	6720
QY	6729	tgcacacttca tctctcaagaagaacgc tgaaccnaa taca ga tga tgc taaagaa tgc	6788
Db	6721	TCCAACACTTATCATCTCTTAAGGAACAGCTGGACCAATATACGATGAGATTAAAGATNG	6780
QY	6789	aaaagctcaagaactat tgaactctgcgzaaaga tga t tcat tga aaaaaaagaa tctcaag	6848
Db	6781	AAAACCTCAGAACTATGAATCTGCGAAAGATGATATCATGAAAAAGAAATCTCAG	6840
QY	6849	tgcctgcgttca tgaagaacagcctca gcaagaaga tga tgc tgaagaac tga tgc	6908
Db	6841	TCGTCGCTTAAATGACAGACAGAGCTCTCAGAGAAAGATATGTTGCGAAGAACTATNG	6900
QY	6909	ataactctgaagacgaaaaccaagaat taca gga gct tga tga aaaaaa tctccgc tgc tgc	6968
Db	6901	AATATCTTTGAAGAGCAAGAAACCAAAATACAGAGCTGATGCAAAAAATCTCCCTCGT	6960
QY	6969	atccaagcaacacacttgc tcaagcag tctctc tga tga gct tcaaaagaac tctgaag	7028
Db	6961	ATTGAGACCAACACACTTGTGTCAGCAGTCTCTCTATGACGTTCAAAAGAAATCGAAG	7020
QY	7029	caacacaaacttga tgttaaata taaagga tctctgtca tccacgc tctcaga tccct	7088
Db	7021	CACACAAACATTGGATGTTTAATTAAGGAATCTGTCATCCACAGCTCTCCACATCTT	7080
QY	7089	tgcgaagcttgc aaactgaagacat ttaag ttaaa tctcaactga ggaagccctcttga aca	7148
Db	7081	TTGCGAGCTTGCAAACAGTACGATTTTAAGTAAATCTCAACTGCGAGACCTTCTTGAAACA	7140
QY	7149	aattaaag tgc tga taccgaactgc tgcag tcaaaagaatca tga tgc tga tcaaaaga t	7208
Db	7141	AATTTAAGTGTGTATACGAACTGCTGCATCAAAAGAAATCATAGCTTGATCAAAAGTT	7200
QY	7209	atggaagaagacctgc tgc tgc tga gaga aagaag gca tga tga gct tgc gaa tccaaactgc gct	7268
Db	7201	ATGAAAGAGACCTTGCTGCTGAGCAAAAGAGGATGATGATGCTGCACATCCAAACGCACT	7260
QY	7269	gttttga gca gca ggcaga aaaa tgc tga tct tgc a tca tca gga gaa c tcaag tgc tgc	7328
Db	7261	GTTTGGACACACAGCGCAAGAAATGTCGGAATTCCTGCATCTGAGGAACCTCAAGTTCTGTG	7320
QY	7329	aaattgaatctct tgaatga tga t tca t ttaaaaa gcaaa taaat tcaag tgc tcaag	7388
Db	7321	AAATGAATTCCTTAAGAGATCTACTTTTAAAAAACCAATTAATTCAGATGTCGCAAG	7380
QY	7389	atgaacttctga gga tgc aagiat tct taa tcaag taa ga tcaaa caa c tgc aagaaga agc	7448
Db	7381	ATGACTTTTCAGAGGTGCAGTATCTCTAATTAAGTAGATTCMACCTGCAAGAAAGAC	7440
QY	7449	ttga gca caa aaga agct tta tgc ag tgc tgc tga gaa at ttc gga gac tgc a gct tga tgc	7508
Db	7441	TTGACACAAAGAAAGCCTTATATGACAGGTGTGAGAGAAATTTGAGATCTGCACTCGATNG	7500
QY	7509	ctaaga aaaa cta ag tga ag gca tgc aaaa aaga aaaa ta ggc gca tgc tct tca cca taa gca	7568
Db	7501	CTTAAAGAACTCAGTGAAGGCATGCACAAAGGAAATATGGCGCATTTGCTTCTTACCATACAGC	7560
QY	7569	tcttaacaaaaagc ttaagc ag tgc t tca g c tgc gaa a taca a g tgc ga taa c g t a t	7628

```
|||||
Db 7561 TCTTAACAAAAAGCTAAAGGAGTGTGTACGTGCAAAATTAACGTAGATTAACCCGTAT 7620
QY 7629 atctgaaccagtttgaagcgaatctgcaagaagaagaacaaacaaagaactatgc 7688
Db 7621 ATCTGAACCAAGTTTGAAGCAATTTGCAAGAGAAAGAAACAAACAAAGACTTATGC 7680
QY 7689 gcsagaatgagcaacacagccctagtgctagtgtatgtgaggaagaatgtctagctt 7748
Db 7681 GCGAATGAGACACCGCCCTAGTGTATGTAGAGAAAGAAATGCTAGACTTT 7740
QY 7749 taggactactgaacacttcaagatgaatccaaagaactcaatcaagatcaaatgc 7808
Db 7741 TAGGCATATGAAAAGTGTCAAGATGATATCAAGAAATCAATCAAGATCAAAATGC 7800
QY 7809 tagaaatgaactgaacttgg tcaagaatgctatgcacaaagtgtgaaaaagttgcga 7868
Db 7801 TAGAAAATGAACGTGAATGTTGTCAAGATGATGCCATGCAAAAGGTGAAAAGTTGCAA 7860
QY 7869 ttctgaagacaaactactaagcagaacgcgcgaagctgaagctaaacgcgaatgcagta 7928
Db 7861 TTTTGCAGACAACTTAAAGCAGAAACCGCAGAGCTGAGTAAACGCAATGCAGGTGA 7920
QY 7929 aactaactaaaagcaagataatctcaagctgcaatgaaagaatagaacactacaga 7988
Db 7921 AACTAACTAAAAAGCAAGATTAATCTTCAGGCTCCAATGAAGAAATAGAAAACCTACAGA 7980
QY 7989 aaatggtgcgaagtgagcgtacccatataaagaagaatlgacaacttaaaactaagg 8048
Db 7981 AAATGTTGGCAAGGTGAGTACATATTAAGAAAGAAATTAACAACTTAAACTTAAGG 8040
QY 8049 tggtaagaattgaatgaaaaataaagtaactcaaaagcaacagcaagaagatgcct 8108
Db 8041 TGTATAAGATTGAATGAAAAATTAAGTACTCAAAAGCAACAGACCAAGATTCCT 8100
QY 8109 acttaagctctgttttgaagaatgaagaagcctgctgaagttaaagaagaactta 8168
Db 8101 ACTTAAGTCTGTTTGGAGATTAAGAAAGAGCTTCGTAAAGAGTAAAGAGACTTA 8160
QY 8169 ggcgagcacagcagacacagatacaacagtttgttccaaaagattatcagaagcct 8228
Db 8161 GCGGAGCACAGGCACACACAGTACACAGTTTGTTCAAAAGATTATCACAAAAGCTT 8220
QY 8229 caacttccctgtgactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 8288
Db 8221 CAACCTTCCCTGTGACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 8280
QY 8289 tgcgtcagtcgaaaaagccgcttgaagaagagcgtgtcaactataagaagaatatac 8348
Db 8281 TGTGTCAAGTCGAAAAAGCCGCTTGGAAGGAGAGCTGTCACTTATTAAGAAATATATC 8340
QY 8349 atcattatacagactatgtcaagttcgaagatcgaagaagaacaaagcaaatctg 8408
Db 8341 ATCATTTATACGAGACATATTCAGATTTCGAAAGATCGAAAGAAACGCAAAATCTG 8400
QY 8409 atgtctatcttctcaactgatacacaagaagcctgcgtctcaaaacttgaactt 8468
Db 8401 ATGCTATTTCTTCTCACTGTGATCATCACAGAGGCTCGCTTCACAAAACCTAAACTT 8460
QY 8469 acagacatgcccctgttactccaagaagctcgaatgcgaagcctcaaacctgaactc 8528
Db 8461 ACAGACATGCCCTGTTACTCCAGAAAGGTGAAATCCCAAGCCTACACTAGATCTC 8520
QY 8529 caaagaagtcggaatccagcactaaagtgctgtgtccaaagaagtgccgaatctaca 8588
Db 8521 CAAGAAGTCGAGTCAGCACTAAACGTGTGTGTCAACCAACAGAGTCCGAATGTACA 8580
QY 8589 gccaatagtaatgtctccagcgaagccggagatgaataacatatacttctccaagca 8648
Db 8581 GCCAATTAGTATGTCTCCAGCAAGACCGGATGCAATAACATATATCTTCCAGACA 8640
QY 8649 aggttgagctgcacaaaagcgtgctgtctcaaacagatgagatgcccacagc 8708
|||||

|||||
Db 8641 AGCTTGACTGCACAAAAAGCGTGCCTGTCTCCAAACAGATGGAGATGCCACCCAGC 8700
QY 8709 atgtcatacccttggaagcagcagcgtgataaatactactgaagcagcttctgc 8768
Db 8701 ATGTATATCCCTTGCGCAAGCCGAGCTGATTAATAATCTAATGAAACAGCTTATTCG 8760
QY 8769 acaattgtctctccatgcaaacagcaaaaagtaacaggaatcaaatctcaatccctaaag 8828
Db 8761 ACAATTTGCTTCTCCATGCAACAGCAAAAAAGTACAGGAATATCAAAATTCCTTAAG 8820
QY 8829 gcaaatatgtgaatgaaatcaaaatcgatgcctactactgtccatctcagtttctga 8888
Db 8821 GCAATATTATTGATGATGAATCAAAATCAATGCTTACTGCTCATCTCAGTTTGTGATA 8880
QY 8889 atttaagcttgatgattttcagagctcaacacagcagaagcgaatgacaaagctcagg 8948
Db 8881 ATTCTAAGCTTGTGTGATTTTTCAGAGCTCACACAGCAGAGCAATGCAAAAAGTACAG 8940
QY 8949 ctgagaactgtgtgtatgagcaaaaaaagaagcgcacctgaaatgtaaaaacatctaga 9008
Db 8941 CTGAGAACTGTGTGTATGAAAGCAAAAAAGAAACGCACTGAATGTAAACATCTAGA 9000
QY 9009 tccctgtacatctgactctcctgtctcgaagaagactgtactctgtctctttagg 9068
Db 9001 TCCCTGTACATCTGACTCTCCTGTCTGCAAAAGACTGTGACTGTGCTTCTGTAGG 9060
QY 9069 aagaacactagaacatgcacatgtctcgtatgaagaagctcactgtgaagcaaaagctgt 9128
Db 9061 AAGAAACACTGAACATCTCCATGTCTGTGATTAAGAAAGTCTCACTGGAAGCAAAAGTTGTT 9120
QY 9129 cttaagtgatcaactgtgtgtgcgagtggttaagctttaaataaagtgcaatgt 9188
Db 9121 CTTAAGTGTATACATCGTGTGCGAGGTGTTACGCTTTAAATAAAGTCAATATCT 9180
QY 9189 ctcaagctattatatagttatgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 9248
Db 9181 CTCAGTCTATTTATATGTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 9240
QY 9249 cataattgtgcatactgttagttcttctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 9308
Db 9241 CATATTTGGCTCATCTGTATGTTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 9300
QY 9309 catltagctgtatgttttactgactgtgtctcctgtgtgtgtgtgtgtgtgtgtgtgtgt 9368
Db 9301 CATTTGAGATGTTTGTTTTACTGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 9360
QY 9369 taagtgtccaaatgttttactactatataatgtgtgtgtgtgtgtgtgtgtgtgtgtgt 9428
Db 9361 TAGTGTCCAAATGTTTTCACACTTATTAATTAATGTGTGTGTGTGTGTGTGTGTGTGTGT 9420
QY 9429 tacaagaatagggcaaaagtgttttctcactaggttaacagatagggttlaactaat 9488
Db 9421 TACAAAAGATAGGCAAAAGTTGTTTCTCACTAGTAAGCAATGAGGTTTAACTTAAT 9480
QY 9489 tatttctcaataaataatctataataatggaatcattataggtgacctaagtgtatcc 9548
Db 9481 TATTTCTCAATTAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 9540
QY 9549 tgcactctgttaaaccttttlaaattcaactcgttaataaagcagccgtatcttaaatc 9608
Db 9541 TGCACTCTTGAACCTTTTAAATTTCACTGTATTAAGCAAGCCCTGATTTAAATTT 9600
QY 9609 aaaaaaaa 9618
Db 9601 AAAAAAAA 9610

RESULT 2
HSCENPE HSCENPE 8257 bp mRNA linear PRI 10-JAN-1993
LOCUS H.sapiens CENP-E mRNA.
DEFINITION H.sapiens CENP-E mRNA.
ACCESSION Z15005
VERSION Z15005.1 GI:29864
```



```

QY 1023 aaattcatgtggaagaaatgctaaacggtataatttcgaactatgcagcttctt 1082
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 968 AGAATTCCTTGGGAGGAATCCAAAGACACGATATATCTGCACATTTCTCCAGTATCTT 1027
QY 1083 ttgatgagactcttaagtactcagttcagttgcagctacccaacatgltgagaatctc 1142
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1028 TTGATGAACACTTCTACTGCTCTCGATTGGCCAGTACTGTAAATATATGAAGAATACTCTC 1087
QY 1143 cccatgttaagtaggtctctgtagttagtgcgtgtgctaaagagtaacgaagaaatct 1202
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1088 CTTATGTAATGAGGTATCAACTGATGTAAGGCTCTCTGTAAGAGGTATGAAAGAAATATA 1147
QY 1203 ttgatllaagaacaataatagaatcttagatgcacgtctgtaacaaagctcaagcaa 1262
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1148 TGGATCTTAAAAACAAATTAAG-----GAGCTTCTTTAGAGCGGGGCTCAGGCA 1201
QY 1263 ttgctaaagaagacatacacagttgtcagctgaatacaacaactacacaaagagagag 1322
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1202 TGGAAAAAGACCAATTGGCCCAACTTTGGAAGAAAAAAGATTGGCTCAGAAAGTACAGA 1261
QY 1323 aagatagatagtgcactgacatgatattgtgtgtg----cttaccccaagaatctaac 1379
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1262 ATGAGAAAATTTGAAAACCTTAACGCGATGCTGGTGACCTCTTCTCCCTCAAGTTGCAAC 1321
QY 1380 aggcacaaagggtcacaacgaacgaagatgacgtggcgccgaagaaatcccaata 1439
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1322 AGGAATTAAAGGCTTAAAGAAAGACGATGTTACTGTGCTTGGCAAAATTAACAAA 1381
QY 1440 gttacatgctctcgtgtgtcttctgacttgatagctatccaagattaccgtgaattta 1499
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1382 TGAAGAA-----CTCAAACTATGCGAGATCAATTAATATACCAACAATATATA 1429
QY 1500 gcaagaagcgaaagttctcagatgccttcattccagaatgtagtaccgtgttcta 1559
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1430 CACAAAACACATPACGTTCTTATTAATTTATTAAGAAATGATGATGCTGT 1489
QY 1560 cagaatttcgatttgatgacgccc 1587
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1490 CAGAGTCTGATGTTTTCAGTAACTCT 1517

RESULT 3
AB028468 2576 bp mRNA linear PLN 25-DEC-1999
LOCUS Arabidopsis thaliana ZCF125 mRNA for kinesin-like protein, complete
DEFINITION cds.
ACCESSION AB028468
VERSION AB028468.1 GI:6526974
KEYWORDS kinesin-like protein.
SOURCE Arabidopsis thaliana (strain:Columbia) cDNA to mRNA.
ORGANISM Arabidopsis thaliana
REFERENCE 1 (bases 1 to 2576)
AUTHORS Kato, A., Suzuki, M., Kuwahara, A., Ooe, H., Higano-Inaba, K. and
Kameda, Y.
TITLE Isolation and analysis of cDNA within a 300 kb arabidopsis thaliana
JOURNAL genomic region located around the 100 map unit of chromosome 1
MEDLINE Gene 239 (2), 309-316 (1999)
20018182
REFERENCE 2 (bases 1 to 2576)
AUTHORS Ooe, H., Kato, A. and Kameda, Y.
TITLE Direct Submission
JOURNAL Submitted (05-JUN-1999) Hidekazu Ooe, Graduate School of
Science, Hokkaido University, Division of Biological Sciences;
Kitaku Kita10 Nishi8, Sapporo 060-0810, Japan
(E-mail:ooe@sci.hokudai.ac.jp, Tel:81-11-706-2743,
Fax:81-11-706-4851)
FEATURES
Location/Qualifiers
source 1..2576
/organism="Arabidopsis thaliana"

```

```

/strain="Columbia"
/db_xref="taxon:3702"
/chromosome="chromosome 1"
/map="between m1303 and m1259"
8..2479
/gene="ZCF125"
8..2479
/gene="ZCF125"
/codon_start=1
/product="kinesin-like protein"
/protein_id="BAA8112.1"
/db_xref="GI:6526975"
/translation="MEKICVAVRVRPAPNGASLWKVEDNRISLHNSLDPPTTASH
AFHVEDSESTNSVYLILKDIHAVEENFGATFAVCGTSSGKFTPTMGSETPDI
IRRSVRPEIRIHSIDREFLIRVSYEINENFNDLAENORLOJHEHGEVFA
GIKEEIVSDAEOLIKLIDSGEVRNHEETNMNYSHSHTIFPMVIESRDKSSDA
IRVSVLWVDLAGSERIAKTGAGGVRLOEBKYLINSLMITGVNINKLSDSTKIRAHIP
YRSKLIRLIQPALGNAKTCIICCTAIPBEHHIEESGTLQFASRAKRLITNCAOVNI
LTDAAILKROKLEIEELRMKLGSHAEVLEDELINLSNOMIKYELBCEBLKQLEBK
RKOEONCIEKDOOMIENLNENFVNSDFRNGSEDFIISRTPDGLCNVNDSDVG
TPCKRSASRSFVARSNNYSGLSDPSFVNSLGVADVEDTMMKLNKEFVLDIOFT
PAYKCOPTPLSIATTECPRENNHSVEDLKSRILDTLNTENSIOVKFMEOVYLSNNIMO
EMSEIKETITVKEIPYRLSESVANCKDYKDYIVTKSLITPKESTPAULLGTEI
TTSILATLETFQSMIDQKTTGSSIDHPLSDHMETLVNKNITTLTLDQAQKDEFL
NSHNGQETALEBKLLKSELIIIEKRYNLEKELCDKOLLEASRESHRLIKEVOF
LKEEROSLDKRIISQSTQRLRVIAASDKENALDLNVEVKKRRKDEEIKHISIAFATRH
KSFSFSEIKSMOKLTTONSKAP"
BASE COUNT 856 a 461 c 589 g 670 t
ORIGIN
Query Match 2.5%; Score 243.2; DB 8; Length 2576;
Best Local Similarity 57.0%; Pred. No. 7.2e-37;
Matches 540; Conservative 0; Mismatches 378; Indels 30; Gaps 4;
QY 272 tcttcaattcgatcggtgttaattcctcagaatcaacagttcaattaccaga 331
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 134 TCTACGCTTTTGATCATGATGTTGACGAGAAATCTACGAATGCTAGTCTACGAGCTT 193
QY 332 atagacagactatcatcagatcagcttgagagatataatgacaaatattgacatc 391
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 194 CTTACTAAGATATCATACATCCGCTTTGAGGCTTTTAATGATGCTACGCTTTGCTTAT 253
QY 392 ggaacagatcttcaagaaagcgtacacaaatgatlgygaacacaaatcattgagcata 451
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 254 GGGGAAGTACTAGTGGGAGAACGCTTACGATGACGGGCTCTGAACATGATCGGAAAT 313
QY 452 ataccacaagccatacagaagttttaaaatttcaagagatcacgaacagaagatt 511
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 314 ATTCGAACATCTGTCGAGATGTGTTGAAAAGATATACATGATATCGATCGAGCTT 373
QY 512 ctcttaagagttcttataatgagatttacaataaactggaagacctaactgtagat 571
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 374 CTCATCCGGGTTTGTGATGGAATTTACACAGAAATTAATGATCTTTTAGCTGTT 433
QY 572 gacagaagaagaagcccttgaatctcgagagatttcaatagaaacgtgtagtgc 631
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 434 GA-----GAACCAAGAGCTCAATTCATGACGATTTGAGCGCTGGAGTGTGTTGCT 487
QY 632 gactgactgagaactgtgtaatggttccctgaacatgtaaacagtgtagtcaaaaaggt 691
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 488 GGTCTTAAAGAAATGTTAGATGCTGAACAGATTCCTAAAGCTTATGATTCGGA 547
QY 692 gaaaacaaacagacattagagagacaaatagatataatagtagtgcacataca 751
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 548 GAAGTTAATAGCACTTTGGCGAGACAAACATGATGCTTCTATAGTAGAGGCTCACACC 607
QY 752 atattgaatgattgttgaagcgcagacagaatgatacccaacaaatlcagagaactgt 811
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 608 ATATTCAAGATGATGATGAGAGCAGGGGAAAGATATAT-----AGTTCT 652
QY 812 gatgagctgcatgatatctcactgaattggtatagatcttgcgtcagtgaaagca 871
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

Db 653 TCAGATGCTATTCGTGCTAGCTTGAATTTGCTGATTTAGCTGATCGCAAT 712
Oy 872 agccaactgagctgaaggtgtgagacttaaggaagctgaacatcaaccgcgttg 931
Db 713 GCTAAACTGGGCTGGTGTGTACGCTTGGCAAGGCAAGCTACATTAATAAGCCTTA 772
Oy 932 ttatcccttgagcaagcttaataagacttagcgacg---ccagcggtgtgattata 988
Db 773 ATGATCTTGGTAATGTCATCAATAAATTAACTGACAGTACAAACCTAAGGCACATATT 832
Oy 989 aactacagagacagcaaacctaccagaattctccaaattcaatgtggagaaatgctaa 1048
Db 833 CCGTACCGAGATGATGACCTGACTGATCTTACGCTGCTGAGTGGTAAATGCAAA 892
Oy 1049 acggttaaatgtgacaaataacgcaag-----ttcttttgatgagacttaagtaca 1102
Db 893 ACTGCAATTAAATGACATGCGACCAAGAGACATCATATTGAGAAATCAAAAGAACT 952
Oy 1103 ctccagtttgcagctacgcacaacatgtgagaatactccccatgttaataaggtcctg 1162
Db 953 CTCGAATTTGCAAGCAGAGCAAGCGCATCACCAACTGTGCTCAAGTGAATGATCTTG 1012
Oy 1163 gatgataagcgtgtgctaaagaaggtacagaagaagaaatcttggattta 1210
Db 1013 ACTGATGCCCTTATTGATGAGCGCAAAAGTTGAGATAGAAGAGCTA 1060

RESULT 4
AB001426 480 bp mRNA linear ROD 19-SEP-1997
DEFINITION Mus musculus mRNA for motor domain of KIF10, partial cds.
ACCESSION AB001426
VERSION AB001426.1 GI:2443267
KEYWORDS motor domain of KIF10.
SOURCE Mus musculus (strain:ICR) 4 week cDNA to mRNA.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 480)
AUTHORS Nakagawa, T.
TITLE Direct Submission
JOURNAL Submitted (25-FEB-1997) Terunaga Nakagawa, University of Tokyo
Hongo, Bunkyo-ku, Tokyo 113, Japan
(E-mail:nakagawa@kinesin.kaiyodai.m.u-tokyo.ac.jp,
Tel:81-3-3812-2111, Fax:81-3-5689-4656)
2 (sites)
REFERENCE Nakagawa, T., Tanaka, Y., Matsuo, E., Kondo, S., Okada, Y., Noda, Y.,
AUTHORS Kanai, Y. and Hirokawa, N.
TITLE Identification and classification of 16 new kinesin superfamily
JOURNAL (KIF) proteins in mouse genome
MEDLINE Proc. Natl. Acad. Sci. U.S.A. 94 (18), 9654-9659 (1997)
FEATURES 97420736
source location/Qualifiers
1..480
/organism="Mus musculus"
/strain="ICR"
/db_xref="taxon:10090"
/dev_stage="4 week"
<1..>480
/codon_start=1
/product="motor domain of KIF10"
/protein_id="BAA22386.1"
/db_xref="GI:2443268"
/translation="IFAAGTASGKTHPMGSEDLGVIPRAIHIDFQRIKAPPEREF
LIRVSEIYNETITDLCAKOKKPLIEDRTRYVYSDLFEVYVTAEMAKMILA
TGEKNRHYGITKNNQSSRSHTIFRMILREKRAEPSCGSAVSVSHIKLIDYDEG

BASE COUNT 162 a 87 c 106 g 125 t
ORIGIN
Query Match 2.2% Score 209.2; DB 10; Length 480;

Best Local Similarity 67.9%; Pred. No. 3.4e-30;
Matches 292; Conservative 0; Mismatches 138; Indels 0; Gaps 0;
Oy 380 aatttgcatacagacagacatctcaggaagaagctacaacatgtggaaccacaat 439
Db 1 ATCTTGGCTACGGCAAACTGCAATCAGMAAGACACACGATGATGCTTCTGAAAT 60
Oy 440 tcatgggcatatgcccccaagccatcacggaagcttcttaaatctacaggaatcacg 499
Db 61 TGTTTGGAGTTATACCCAGGGCAATTCTGACATTTTCCAAAGATAAAGATTTCCT 120
Oy 500 aacagagatctctcctaagaagttcttataatgagatcttaacatgaaactgaaagac 559
Db 121 GAGAGAGATTCTCTCAGAGTTCTTAACATGAGAGATCTAATAAGAAACCTTACAGAT 180
Oy 560 ctactgtgtatgacagaagaagacccttggaaatctcgagagatttaataagaac 619
Db 181 TTACTCTGTAAATGCAAAAATGAACCTTGTGATTAATTCGGGAAGACACCAATGAGACT 240
Oy 620 gtgtatgttgcctgacactgacgaagaactgtatggttccgaaacatgtatcacgtgg 679
Db 241 GTATATGTTTCTGATCTCACAGAGAAGGTATTATACAGAGAAATGCTTTGAATGG 300
Oy 680 alcaaaagggtgaaaaaacaagacatataggaagactaaatgaaatcatagtagt 739
Db 301 CTCGCAACAGAGGAAAAAAGACACATTATGSAATACCAAAATGATCAAAAGACAGT 360
Oy 740 cgttacaataatcttagaatgattgttgaagcggagacagaagaatgccacaat 799
Db 361 CATTCTCATACATTTTATGATGATCTGMAATAGAGAAAGCTGAGCCTTCTAAC 420
Oy 800 tcaaggaact 809
Db 421 TCGATGGCT 430

RESULT 5
AX321334 581 bp DNA linear PAT 15-DEC-2001
LOCUS AX321334
DEFINITION Sequence 351 from Patent WO0177168.
ACCESSION AX321334
VERSION AX321334.1 GI:17905139
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (sites)
AUTHORS Lodes, M.J., Wang, T., Mohamath, R. and Indrias, C.Y.
TITLE Compositions and methods for the therapy and diagnosis of lung
JOURNAL cancer
MEDLINE Patent: WO 0177168-A 351 18-OCT-2001;
FEATURES CORIXA CORPORATION (US)
source location/Qualifiers
1..581
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 205 a 101 c 130 g 145 t
ORIGIN
Query Match 2.1%; Score 206.6; DB 6; Length 581;
Best Local Similarity 68.4%; Pred. No. 1.1e-29;
Matches 333; Conservative 0; Mismatches 144; Indels 10; Gaps 3;
Oy 955 gaagcttgcagcagcagcgtgtggtgattataactacagagacgacaactcacg 1014
Db 72 GAACTTGTGATGACAGAGTTGGTGTTCATTAATTAATGACAGTATACAGTTAACG 131
Oy 1015 aattctcaaaatcattgtggaggaatgctaaacggttaataatgtgacaaatcacgc 1074
Db 132 AATTCACAGATTCTCTGGAGGAAATGCAAAAGACAGTATTATCTCACAAATATCTCC 191

Query Match	Best Local Similarity	Score	DB 3:	Length	7524:
Matches 464; Conservative 0; Mismatches 387; Indels 24; Gaps 3;	53.08;	177.8;			
Db 360	tcgaaggaataaataatgacaaatatttgacaaagacagacatcttcaggaagaagctlaca	419			
Db 548	tccagagccttCACGCGAAGCTATTTTTCCTACGCGCAGACGTCGTCGGGCAAGAGCTACA	607			
QY 420	caatgaatggaacacaaatcaltcatitggcgaataaccccaagcgaataagaaatlttta	419			
Db 608	ccatgatggcgagacagacagaaatccggcgctatagtgctaacccccaaggaagatcttcc	667			
QY 480	aaattattcaggaataaccgaaagaagagtlctcttcataagagttcttataatgagaattc	539			
Db 668	AACGAGATTCCTCAGTGAAGACGAGCGGAGCTTCTGCTCGCGTGGGCTACATGCAAGATCT	727			
QY 540	acaatgaataactgtaagaagactactctgtagtgaagaagaagaagaagcccttggaattc	599			
Db 728	ACAAGAGGAAGATCTACGATCTGCTG-----AACCAAGAAAGATCAGACCTCAAGATCC	781			
QY 600	ggcaggaatttaataagaacgttgatgtctgacgtcagacgtgaagaaactgtaagtcttc	659			
Db 782	ATGAGTCGCGGAATGAGATGGAATGT---GAACCTGGAGAGATGCATATATACACAGCG	838			
QY 660	ctgaacatgtaataacagtgatcaaaaaaggtgaaaaaaacagacattatgagaagacta	719			
Db 839	AGGTTGACCTCTCCGCCCTGCTATCTTTGGGCAATTAAGAGAGCGCACCGCTGGGAGACGA	898			
QY 720	aaatgaatgatcatagtagctgtcacataacattatgaagaatgatttttgaagaccccca	770			

Db 899 ACATGAACGACGCTCAGCCCTTCGATGTCATATTTAAGATATCATGAGCTCCGAA 958
 Qy 780 acagaataatccccacaattcagagaactgtgagtgcgtcatgtatctcacttga 839
 Db 959 AGTCGA-----CCACAGCATGACACGACGAGTGAATCAAGCGCTGTGA 1003
 Qy 840 attgttagatctctgtcgtcagtgatgaagcaagcccaactgtgagctgtgagac 899
 Db 1004 ACTGTGATTTTGGCTGGCTGACAGCCGGCGGACCAACGGGTGCTCGAGGAGCGGCC 1063
 Qy 900 ttaagaagagctgcacaacacccagctgtgttacccttgacaggtatataagaagc 959
 Db 1064 TCAAGAGGGCGGTCTATTAACAGAGCTTACTTTTCTCATACGTTATCAAAAGTC 1123
 Qy 960 ttagcagcggcagcgtcgtgtgattataaactacagagaagcaaacccaagaattc 1019
 Db 1124 TGTGGGAAACCGGACATAGATTATTAACACACGCGCATTCACAGCTGACCGCATTC 1183
 Qy 1020 tccaaatcatcttgaggagaatgtcctaaaggtataatttcacaaattagccagttc 1079
 Db 1184 TCGAGGCTTCTTGGCGGCAATGATTCATCATTCATTCACAAATCAACATCGA 1243
 Qy 1080 ctcttgatgagactcgaagtaactcagcttgctgacgtacgtcgaacacatgtgagaata 1139
 Db 1244 TCATGGAAGAGTGCAGCTTACCTTACTTGTCCACGCGGACCAAGAAATTCGCATTA 1303
 Qy 1140 ctcoccatgttaatgagtcctgtgagtgaagcgcttgctaaagaagtcagaagaaga 1199
 Db 1304 ACCCCCAAGTCAACGAGATGATCTGACGCCACATGATGAAGCGCTGAGACGAGAAA 1363
 Qy 1200 tcttgatttaagaagaacatagagaattagag 1234
 Db 1364 TAAAGGTCTTAAGATAGCTGGCCGGAAGAAGAG 1398

RESULT 7
 AB062739 3692 bp mRNA linear PIN 23-NOV-2001
 LOCUS Arabidopsis thaliana MKRP2 mRNA for kinesin-related protein,
 DEFINITION complete cds.
 ACCESSION AB062739
 VERSION AB062739.1 GI:16902293
 KEYWORDS Arabidopsis thaliana cDNA to mRNA.
 SOURCE Arabidopsis thaliana
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (sites)
 Itoh, R., Fujikawa, M. and Yoshida, S.
 Kinesin-related proteins with a mitochondrial targeting signal
 Plant Physiol. 127 (3), 724-736 (2001)
 2 (bases 1 to 3692)
 Direct Submission
 Submitted (04-JUN-2001) Ryuichi Itoh, RIKEN, Plant Functions
 Laboratory; 2-1 Hirosawa, Wako, Saitama 351-0198, Japan
 (E-mail:ryuichi@postman.riken.go.jp, Tel:81-48-467-9525,
 Fax:81-48-462-4674)
 Location/Qualifiers
 1..3692
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 301..3468
 /gene="MKRP2"
 301..3468
 /gene="MKRP2"
 /codon_start=1
 /product="kinesin-related protein"
 /protein_id="BAB71852.1"
 /db_xref="GI:16902294"
 /translation="MASSSSRRSSRPPSPASSSTSSHLNLRIPRNSYSTASSLITS

AAGIASRMSPTSFSDSGLIGSGFSGSPVYPVSEELLGPMDDTISSEBDSISVT
 VRRPLSDREYORGDENVAMYPDGDTLLVREHYPLTAANDKXVEGQATTDIVDVAAR
 PVYKAAMEBENVTFAVGYTSSGKTHRMHGDDESGLITLAKDVFSLITQDPGPEPL
 LRVSTLEIYNEVINDLDPGONLRVEDSGTYVEGKEEVVLSPLHALSTIAGEE
 HRHVGSNNFNLLSSRHTFTLVMSVSAHDYDDVTFSQLMLDLAGSESSKFTTG
 LRKREGSYINKSLTLTGITGLKESKATHDYRDSKLTRLRLQSLSGHVSILCTI
 TPASSSEETHNTLKFAPRAKSIETIYASNOIIDEKSLKTKRISTKLIDDLOR
 GMVGVSHHEIMSLKQOLEEGOVKMSRLFEEDKAKLMSTRIKRLTLIVSTKNSI
 PGYSGDIPTHQSLNSAGKDKFDSLSEBNDIGSPSTLALISEGSLGRNHRSSKL
 NDENSPGAETFGVMPDEIDLVEQVKMLAGEIAFVSTLRLVDQSVNDENSOQO
 IONLREIHEKOROMHGLEOLIIEGSEASIANSLVEMOKVMSLMTQSNESFLEI
 KSDANCIOLOEKTEENKLEHERYNLEORLNVSESKSPSSCNKASVSEVDEL
 KKIOSEIENEELKEHVOIVEENSGALHVOOKLAEEVSKYKELASAAVLEKLILAS
 EYTKLSIOWTKLEKELAAARDLAOTRNPNVNGNRRKNCAEBSRGRTSSRSDDER
 DAMNLDPEDLMELVOROREVALSALAEKFTIDETRKKAEBARKREALELNLAN
 MNVLAKLKKDNGALPEPNTDPRGRLELESQHAVALKEQVSSAPRQPPVVAAETEE
 TPKEBFLVARLKRMOEMKEKEMKSGANDANSHICKVCFESPTAAILILPCRHPLCK
 SCSLASECPICRTKISDRLFAFPS"
 BASE COUNT 1160 a 740 c 842 g 950 t
 ORIGIN
 Query Match 1.8%; Score 176.4; DB 8; Length 3692;
 Best Local Similarity 54.6%; Pred. No. 5,7e-24;
 Matches 523; Conservative 0; Mismatches 396; Indels 39; Gaps 7;
 Qy 273 ctccaattcgatcggtatcttaattcocaagaatcaaaagtcgaattaccagaaga 332
 Db 713 CTTCACCATTTGATTAAGATTGTTGGACACAGCACTACTATTGACGTTATGATGATG 772
 Qy 333 tagcaatcactatcagatcagctcgttcgagagatcatalgcaaaatattgtcaga 392
 Db 773 CTGCGAGACCATGATGTCAGCGCAATGGAAGCTGTTAAGGACCGTTTTCCTTACG 832
 Qy 393 gacagacatcttcagagaagcgtacacaatgtatggagaacacccaattcatttgycata 452
 Db 833 GTGTTTACTACGATGGAAGACACATACATGATGATGATCAAGANATCTCCGGGATCA 892
 Qy 453 taccacaagccatcacagaagtttttaaatltaacagagatcaccagaagaagtttc 512
 Db 893 TACCGCTACCAATTAAGAGACGCTGTCATATCATCCAGATACTCCAGAGGAGGACTTC 952
 Qy 513 tcttaagattcttatalagaggaatttacaatgaatcgtgaagaagcctactcgtgag 572
 Db 953 TGCCTCGGTTCATATCTTGAATATACATGAGCGTGAATATGATTTCG-----G 1006
 Qy 573 acagaagaagaagcccttggaattcgcgaggaatttaataagaacgtgatctgcg 632
 Db 1007 ATCCAAACAGGCCAAACTTACGCTGTAGAGAGGATTCACAGG---CACTTATGTTGAG 1063
 Qy 633 acctgaactgaagaactgttaattggtccgtgaacatglaaatacagtgatcaaaaaggtg 692
 Db 1064 GTATCAAGAGCAAGAGTGTGTTTGTCTCGGCCATCAGATCATTCATTTGACGCTGGG 1123
 Qy 693 aaaaaaacagacattatgagagacccaataatgaatgaatcattcgtccatcaaa 752
 Db 1124 AAGAACATGCTGATGTTGGTTCAATATTTTCAATTTGTTGACACACAGACATCAACATTA 1183
 Qy 753 talttagaatgattgttgaagaagccagagacagaatgatccacaataatcagaagaactgtg 812
 Db 1184 TATTTACACTGATGTTGAAGATG-----CTACTGGAGAGTGAATATG 1228
 Qy 813 atgagagctgcattgatactcactgtgaattcgttagatctcgtcagtgaaagaaga 872
 Db 1229 ATGGA---GTATCTTTTCTCAACTCAATTTGATTGACTTAGCTGATCTGACAGTTCCA 1285
 Qy 873 gccaacttgagctgaaagtgtagagcttaagaagaagccgaacacacacacagcagcttgc 932
 Db 1286 ---AAACTGAACAACTGATGATGAGAGAGAGGCTTCATACATCAACAAAGAGCTTTC 1342
 Qy 933 ttaacttgacaggtatataagaagccttagcagcagcagcgtcgtgtgattatataact 992
 Db 1343 TAACTGTGGAAGTGTGATTTGGAAGAACTTAGTGAAGGTAAAGCACTACATTCCA---T 1399

RESULT	8
LOCUS	AB062738
DEFINITION	AB062738 3273 bp mRNA linear, PLN 23-NOV-2001
ACCESSION	Arabidopsis thaliana MKRPI mRNA for kinesin-related protein,
VERSION	AB062738
KEYWORDS	AB062738.1 GI:16902291
SOURCE	Arabidopsis thaliana cDNA to mRNA.
ORGANISM	Arabidopsis thaliana
REFERENCE	Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
TITLE	Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
JOURNAL	1 (sites)
MEDLINE	Itoh, R., Fujiwara, M. and Yoshida, S.
REFERENCE	kinesin-related proteins with a mitochondrial targeting signal
AUTHORS	Plant Physiol. 127 (3), 724-726 (2001)
TITLE	2 (bases 1 to 3273)
Direct Submission	Itoh, R.
Submitted (04-JUN-2001)	PubMed: 11461446

FEATURES	Location/Qualifiers
source	1..3273
gene	/organism="Arabidopsis thaliana"
CDS	/db_xref="taxon:3702"
	259..2931
	/gene="MKRPI"
	259..2931
	/gene="MKRPI"
	/product="l
	/codon_start=1
	/protein_id="BAB71851.1"
	/db_xref="gi:16902292"
	/translation="MSARSGRSSTIPANPRSPATIPMKRPETPSSSHFSAAPVNS SYLIRLSSPSSPTSSAAASTVAATSTKLKNIVITIREPLSPRNGNDEITAYAG DYKTINENPISLCTGFDFVREPPPTTRKIVYDIITREPLSPRNGNDEITAYAG KTHHMGESRSPGIIPLAVKDVSIIQETPERPELFRVSYLETVEIVYENIDLDPTGG LIRHDSOGTVEGIDKDEYVSPAHASIIASEEHRHVGSSNNVLSNSHMEFTLT IESPHGKGDDGEVNSLSOLHLDIAGSSSKETETGCRKRGSSINRSILTGTIVS KLTDKKAHPIYRDSKLTRLTLOSLSGHRVSLCTITPASSTSEETHNLTKEAQRK HEIKASRNKIMDEKSLIKKQKEICSLDEELITOLHGNODLADRLKLYVLTSLRED DEEAKALMGRITORTYKTLITVYSTKSLQAAVDPHIMROAFGEDELAIVPDRRRRM ADGASVSTSEHLKEPRDGNSSLDEMTKDRANKNRGMILKTLKSDGAVGLPTG NOSASGSPSSSSKYTOTKTTTRRENNAAIKSPETKTAGDGLFSAVTPEDSPSTGTTI ADOMDLHEQRTKLVEVALRTLSLNRKLEQAAARPEDFHROIOKLEDEISEKQD IRVLEQOIIETFGMTPYASDGLGMPVLSLQMOLEKTFHEHLSKSDNITLQEOLOM TKRSNMAEQETIILRQOLDSLAERSTQOLAGDSSSKNHNNGEESITVSGAGTP TSVSLNVRFAOETKEIYNETALNSQALEINLKNKMKRLIEKKELGKLKNTKTEE ASYAKETLASAAVEIOLNAEYTRILCNENAKISR"
BASE COUNT	648 c 708 g 863 t

Query Match	Best Local Similarity	Score	ID	Length
1.7%;	53.2%;	167.6;	DB 8;	3273;
Matches 509;	Conservative 0;	Mismatches 414;	Indels 33;	Gaps 6;
QY	281	ttgatcgtgtatattcaatctccagcaatcaaaagtaaatcttcccaagaatagcaagta	340	
Db	610	tttgatagagatatttgctccacacacactacgtgctgtttttgacatgtgctgtca	669	
QY	341	cccatcaatcagatccagcttgcaggaatataatgagcaaatatttgatacagacaga	400	
Db	670	caagttgttagtgaggcgaatgacagagatgtaagtgactgtgttgattgagagttact	729	
QY	401	tcttcagcaagcagtlacaaatgtagggaaacacaaatctatgtggcataatccca	460	
Db	730	agcagtgagggaacacatcatatgcatggagacgaagctacaccttgaaatgctgtg	789	
QY	461	gcatcaggaaggttttaaaatattccagagagatccgagagagatcttctctaaga	520	
Db	790	gctgtgaaagatgtattttgacatttttcagagacacacgaacgggaatttcttctgt	849	
QY	521	gttcttataatgagattacaagaacatgtaagactctgtgtgagacagaaga	580	
Db	850	gttttgatcttagagatattacatgacgtcatatgactttact-----agatccaca	903	
QY	581	aagaagccctggaaatctgcagagatlttaataagaagctgtatgtctgactgact	640	
Db	904	ggcgaataatcttmacatgacagagagattctcagg--cacatattgtgaaggtatmaa	960	
QY	641	gaagaactgtlaatgcttctgacacatgtaaaagtgatcaaaaagtgaaaaaac	700	
Db	961	gatagagttgctttatccctgcctcactgtctatcccttattagacatctggagaagac	1020	
QY	701	agacatattgaaagcactaaatgaatagatagtagtgcgttcacatacatattaga	760	
Db	1021	agcattgtttgttcaaatatgactcatctttttacacatctcgagccacacatgtttcac	1080	
QY	761	atgattgttgaagccgagacagaatgatcccaaatltaaggaactgtltagact	820	
Db	1081	ttgactattgaaagcag-----ttcacatggaaaggaagtgacggcgaaagat	1128	
QY	821	gtcatgtgttcacttgaatttggtagacttgcgtgacgtgaaagagcaagccaaat	880	
Db	1129	gttagctgttccaaatgacaccttattgatcttctggtctgtagaagttctga---aaact	1185	
QY	881	ggaagctgaagtgtagactlaaggaagcgtgaaactcaacccgacgctgttactct	940	
Db	1186	gaaatgaaactgacacgaaagaaagagagccttcttcttcaatgaaaggtttactcact	1245	
QY	941	ggaagcgttattaaagacttagcgagcgcagagctgtgtgattttaaaccacagagac	1000	
Db	1246	ggcactgttattatcttaattgacagacacacaaagcagct---catatgacctacaggaat	1302	
QY	1001	agcaactcacacgaattctccaaatcttctgtgtgaggaagatgctaaagcgttaaat	1060	
Db	1303	tcttaacttttagcgctcttgcagctacacactgaagggctacgagcagcttactatt	1362	
QY	1061	tgacaatttgcagcttctt-----ttgatgtgaacttaagtaactcagttgtgc	1114	
Db	1363	tgcactgtacacacctgcatcacaacacatgaaagagcgaacacacactgaattgct	1422	
QY	1115	agtaactgcaaaatgtgaaatactcccatgttataatgtgtcctgagatgaaagc	1174	
Db	1423	cagagtgacgaagcatgttgaaatcamaagctttacgaataatgattatgtgagagctc	1482	
QY	1175	ttgtctaaagtgacagaagaaactgtgattttaaagaacaaataggaattt	1230	
Db	1483	cttattataaaagtatcgaamaagaaatattcgttttacaagaagaaactttagcattt	1538	

LOCUS	AV039966	3599 bp	mRNA	linear	PLN 24-JUN-2001
DEFINITION	Arabidopsis thaliana putative kinase heavy chain (F3K23.14/At2g21380) mRNA, complete cds.				
ACCESSION	AY039966				
VERSION	AY039966.1				
KEYWORDS	GI:14532683				
SOURCE	F11 CDNA.				
ORGANISM	thale cress.				
REFERENCE	Arabidopsis thaliana Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta: Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae: eucroids II: Brassicales: Brassicaceae: Arabidopsis. 1 (bases 1 to 3599)				
AUTHORS	Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K., Dale,J.M., Gibson,H.A., Goldsmith,A.D., Jiang,P.X., Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Koeseema,E., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Sakou,M., Seki,M., Shinn,P., Southwick,A., Tracy,S.E., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.				
TITLE	Full Length cDNA of gene F3K23.14/At2g21380 (GI:4567271)				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 3599)				
AUTHORS	Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K., Dale,J.M., Gibson,H.A., Goldsmith,A.D., Jiang,P.X., Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Koeseema,E., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Sakou,M., Seki,M., Shinn,P., Southwick,A., Tracy,S.E., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.				
TITLE	Direct Submission				
JOURNAL	Submitted (11-JUN-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA				
COMMENT	RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA : 'RIKEN Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.				
<p>The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAPL cDNAs: Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K., Dale,J.M., Gibson,H.A., Goldsmith,A.D., Jiang,P.X., Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Yu,G., Bowser,L., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Koeseema,E., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Tracy,S.E., Davis,R.W., Ecker,J.R. and Theologis,A.</p> <p>Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC) contributed equally to this work as PIs.</p>					
FEATURES	location/Qualifiers				
SOURCE	1. 3599				
/organism="Arabidopsis thaliana"					
/db_xref="taxon:3702"					
/chromosome="2"					
/clone="RAFLO5-09-N13 (R10023)"					
/note="This clone is in a modified pBluescript vector (lambda ZAP) as a XhoI/SstI insert.					
ecotype: Columbia"					
1. 3599					
/gene="F3K23.14/At2g21380"					
1. 119					
/gene="F3K23.14/At2g21380"					
120. 3296					
/gene="F3K23.14/At2g21380"					
/codon_start=1					
/evidence=experimental					
/product="putative kinesin heavy chain"					
/protein_id="AAK64143.1"					
/db_xref="GI:14532684"					
/translation="MASSSSRRTRSSPFSHRPPSPYSSASTSSLLNNRLLPRSSSS					

Qy	273	cttccaattcgcagtcgtatcttcaattctcagcaatccaacgaagtcgaattcccaagaa	332
Db	550	CTTACGAGTATGGATTAAGTTTGGACACACGTAACACCTCAGAAATATATGACGTAG	609
Qy	333	tagcagtaacctatatacagatcagcttgcaggaatataatgcaacaatattgcatacg	392
Db	610	CTGCTAANACCTGTAGTCAAGCGTGTATGGAAGGTGTAAAGCAACCGTTTTCCTATG	669
Qy	393	gacagacatcttccagcaagcgtacacaaatgcatgggaacaccaaatcattcgggacaa	452
Db	670	GTTGTTACAGACAGTGAAGAACACACATACATGATGATGATCAAGATTTCCTGGATCA	729
Qy	453	taccccaagcacaacgaagattttaaataattcagaagataaccgaacgaagagcttc	512
Db	730	TACATTACCAATATTAAGACGTGTCTGATTCATTTAGGAAACACCGGACGAGAGTTTC	789
Qy	513	ttccaagaagttctatataagagattacaatgaagaactctgaaagaccctactctgacg	572
Db	790	TGCTTCGTGTTTCATATCTTGGAAATATACATATAGGTGATTAATGACTTACTG	843
Qy	573	acagaagaagaagacccttggaaatctgcagaggaatttcaatagaagaactgtatcgtcgt	632
Db	844	ATCCACACAGGCAAAATTTACGTATTAGAGAGATTCCACAGG--CACTTACGTTGAGG	900
Qy	633	acctgactgaagaactgtataatggtctccgtgaacatgtaatacagatgatacaaaaaggtg	692
Db	901	GTAACAAGAGGAAGTGTGTTTGTTCACAGCCCATCGTTATCTTTCATTGACAGCTGGCG	960
Qy	693	aaaaaagacacatataagagagagactaaatagaatgatacctatagtagtctgtaacaataa	752
Db	961	AAGAACATCGCATGTTGTTCATTAATTAATTTCAACCTGATGACACGACGACAACTCACANA	1020
Qy	753	tattagaagatgttctgaagccgagacagaaatgatacccaacaaatccaagaagactgt	812
Db	1021	TATTACACAGTATGATTAAGAAAGTAGT-----CTCATGAGACCAATATAG	1065
Qy	813	atggaagctgatacgtatcctcaactgaattcgttagatcctcgtcggcagtgaaagacaa	872
Db	1066	ATGGA---GTTATTTTTCACACTAAATTTGATTGACTTAGCTGATCTGACAGTTC--	1120
Qy	873	gccaaactcgaagcgtgaaggtgtgagacttaagaagaagcgcgaacatcaacccgacagctgtc	932


```

/gene="SPKR85"
/codon_start=1
/product="SPKINESIN-II (KRP85/95) 85kd subunit"
/protein_id="AA16098.1"
/db_xref="GI:295246"

/translation="MPGSSGNDNVRYVVCNRPUNSKETQGRKSVKMDENKGVYV
INPAPSGEPKSTFEDTVPAPGAKQTDVYNQTAIVDA1IEGYNTIIPAYQQTG
KTFMEVRSQPELRLGIIPNSFAHIEGHIAKEQVNFYLRVSYLEIENKVDLIG
DOQHRLEKRPDPVYVYKDLAFVNNADMDRIMTIGKNSVGTAKNNESRSRH
AIFITLERSDMGIDKEOHYRVGLHMDVLAGSROTKGTGQORKEATKINLSST
LGNVYSLVDGKSTHPIRYNSKILRLIDLSGGNAKVMGANTGAPAEYNDERTSLR
YANAKYIKKAKINEDPKALLREFKELEBELKQISEGEGIDDESGESGDE
EAGEGVKKRRKKNRPKLSPEITMAAMOKIDBEKVALEKEDMEEDNTVHRELQ
RSESELHKAQDDQKILNEKLNALQKTLIVGVVDLASEQDLDEQSALEMERAK
QESMRKMEEREDQRMIDIEKYSLODEAKGRTKLKVTMTLMOAKSEVADQAEHQ
REMEALLENRELSRELRLSMLIIDSFIPQEPQEMIEQYVHNMDIEGMOCLKVAATG
NNMRKQCPVADKDKSLAYGEADLSNVFLYINLEGGGKMYPSQGSRRPRTSSGRPT
GKKKOAAMASIDALLO"

BASE COUNT      722 a      453 c      625 g      413 t
ORIGIN

Query Match      1.6%; Score 153.2; DB 3; Length 2213;
Best Local Similarity 52.2%; Pred. No. 1.9e-19;
Matches 533; Conservative 0; Mismatches 448; Indels 41; Gaps 7;

OY 231 agcgtgaaacacacacattcccaagltgatyggacaaglttcattcagtcgtg 290
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 176 AGGTACACAAACCCCAATGACACCCAGCGGGAGCCCAAGATCGTTACACAG 235

OY 291 tatttaattccacgaatcaacagltcaattaccagaagaatgacgtactacac 350
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 236 TGTTCGCGCCGTCAGACGACACTATGCTACCAACCAACGCGGACGACCTGCTAG 295

OY 351 gatcagcttcgaggaataatgacaataattgcatcagaacatcattcagca 410
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 296 ATGCTATCATTTGAGGGGTACACAGGACTATATTTGCTACGCTACACAGTACAGAA 355

OY 411 agacgtacacaatgatyggg-----acaccaattcattgagcataataccccaag 461
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 356 AGAGCTTACACCATGGAAGAGTACATCTCAACGAGAACTCAGGGGATCATTTCCCAACT 415

OY 462 ccatacagggaagttt-----aaatatccagggagataccgaaagagattctctaa 518
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 416 CATTGCTCACATTTTGGACACATGACTTAAAGAGCAAGAAATGTCAGGTTCCCTACTTA 475

OY 519 gagttcttataagagatttacaatgaacctggaagacctactgtygatgacagaa 578
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 476 GCGTGTATATTTGGACATCTACAAATGAAAGAGATTTTATGCGCAAGAAC--C 532

OY 579 gaaagaaacctggaataatcgcgagatttaatagaacgtygtactgacgtga 638
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 533 AGCAGCATAGATTTAGAGGTCAAAAGCGCGGATGTTGGTGTCTATGTCAAGACCTTT 592

OY 639 ctgaagaacttgtaatggttcctcctgaacatgtaatacagtygatcaaaaagtgaaaaa 698
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 593 CTGATTGCTGTCAATTAACGACAGATGATATGACAGAGTAATGACTCTAGGAAACAAA 652

OY 699 acagacatltgagagaaactaaatgatatcatagtagtcttcaatcacatatatta 758
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 653 ACAGATGTGTGGGGCGACAAACATGAAATCAATCAAGTTCTCGTTCCCAATTTTCA 712

OY 759 gaatgattgttgaagcgcgagacaagaatgataccacaanaatlcagaagaactgtgatgag 818
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 713 CCATTAGCTTGGAAAGAGTACATGGGT-----CTGGATTAAGAACAGC 757

OY 819 ctgtcataggtatctcaacttgaatttgtagatcttctgycagtgaaagaagcca 878
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 758 ATGTCTACAGGTGGGAGAGCTCACAATGTTGATCTTCCGAGAGCTTCCGAGAGACACAGA 817

OY 879 ctgagactgagagtgtygagaaagcgtgcaacatcaacccgactgtttatcc 938
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 818 CCGGTGCGACCGGACAAACGCTCTCAAGGAGGCGCAACAGATCACTGTCTCTTACCC 877

```

```

OY 939 ttggacagttattaaagaagcttagcgacgcccagcgtggtgattataactacagag 998
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 878 TGGTATACGTCATCTCATCTCTGTTGATGTGTAAAGACTAC--ATCCCCTACAGGA 934

OY 999 acagcaactcaccagaattctccaaatcattcattgagaggaatgctcaaacggtataa 1058
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 935 ACTCAAGCTCACCAAGCGCTTACAGACGACGCCCTTTGAGAGCAACGCAAGACAGTCATGT 994

OY 1059 ttgcaacaattacgcagct-----tctcttgatgagactcgaagaactcagttg 1112
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 995 GTGCCAACATTGGACGACGACAGTATGATGATGAGAACATTTAGCACACTAGATATG 1054

OY 1113 ccagttactgcacaacatgtagaatactcccaatgtaatgagtgctctgtagatgaag 1172
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1055 CAAMCCAGCGCAAGAACATCAAAAATTAAGCAAAAGATCATGAGATCCCAAGAGATCCC 1114

OY 1173 cgttgctaaaaagtgtagaagaagaaactcttgatttaagaagaactagagaattag 1232
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1115 TCTTGT--AGAGAGTTCCAGAAAGAAATTGAAGATTCAGAACAGATCTGAGATCTGG 1172

OY 1233 ag 1234
    ||
Db 1173 AG 1174

RESULT 14
AB072777
LOCUS
DEFINITION
Macaca fascicularis testis cDNA clone:Qtsa-20515, full insert
sequence.
ACCESSION
AB072777
VERSION
AB072777.1 GI:16041157
KEYWORDS
oligo capping; f1s (full insert sequence).
SOURCE
Macaca fascicularis adult male testis cDNA to mRNA.
clone_1lb:macaque testis cDNA library Qtsa clone:Qtsa-20515.
ORGANISM
Macaca fascicularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
Cercopitheinae; Macaca.
1 (bases 1 to 2759)
AUTHORS
Hashimoto,K., Osada,N., Hida,M., Kusuda,J., Tanuma,R., Hirai,M.,
Tero,K. and Sugano,S.
TITLE
Isolation of novel full-length cDNA clones from macaque testis cDNA
libraries
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 2759)
AUTHORS
Hashimoto,K., Osada,N., Hida,M., Kusuda,J. and Sugano,S.
DIRECT SUBMISSION
Submitted (09-OCT-2001) Katsuyuki Hashimoto, National Institute of
Infectious Diseases, Division of Genetic Resources; 23-1, Toyama
1-chome, Shinjuku-ku, Tokyo 162-8640, Japan
(E-mail: khashi@nih.go.jp, URL: http://www.nih.go.jp/yoken/genebank/,
Tel:81-3-5285-1111(ex.2120), Fax:81-3-5285-1181)
COMMENT
Lab host: TOP10
Vector: pME18S-FL3 (Acc.No. AB009864)
R. Site1: DraIII (CACTGTCG)
R. Site2: DraIII (CACCAGTGTG)
Description: 1st strand cDNA was primed with an oligo(dT) primer
[ATGTGCGCTTTTCTTTTCTTTTCTTTT]; double-stranded cDNA was synthesized
using specific 5' and 3' primers and amplified by PCR. The PCR
product was digested with SfiI and size selection PCR product was performed to
exclude fragments <1.5kb. The SfiI-digested PCR product was cloned
into distinct DraIII sites of pME18S-FL3. XhoI sites just outside
the DraIII sites can be used to isolate the cDNA insert. Libraries
were constructed by oligo-capping method
(Sugano et al., University of Tokyo, Institute of Medical Science).
Custom primer used for sequencing
(5' end primer [CTTCTGCTCAAAAGCTGCG];
3' end primer [GCAGCTGCGTCAAGCAACA]).
FEATURES
source
1..2759
/organism="Macaca fascicularis"

```



```

/translation="NSKSSSESVRVVHCRPMNSKELAGYREVVDVKLGQVSVK
VKGATNELSKTFTPATYDSDNSKQVEIVEMPREPLVDVILGLNCTIRATGOTGTX
TYMEGRDDEPERKGVIPNSFEHTFHHSRQNOUOLVRSYIEIYOEIRDLISDO
SKRLERKPERDIGNVVKDLSSFEVTKSVKLEHVMNMGONRSYAGNMDHSHSRNAI
FMITIECSQIGLDEHNIIRYKLNLDVLAGSEKOTGAGEELKATKATINISLALG
NVSALVDGSRSTHPIYRDSKRLRLDLSLGNKTYVMVANIIGASVNEVETLTLTLYS
NRAKNINKNPRVNEDEKDALREFOEIAIRKLRVSKRRRGRRTAGGEVDE
DEDEBDLYEGDDVLKMDKEDYREOERIEIEKKLLEDHSLVAERKLLKKEKEM
DLKREAMEMLNSKYKAMESKLINGKRYVHTNEOOKLLEOKRHEIAEOKCRERE
MOOQESODEENLEKETYSSLOOEVDIKTKLKLFLSKQAVKAEIHDIOEHITER
QELFOQNELTRELKHLIENFIPLKENKMNRCFEDEEDSKMLKALISLDMQO
MMRRPVSAIGNRPLCOHAKSMVMRPSRYRAENIVLELDMPSRTTRDYESPALAP
KVQAALEAALQDDDDIQVDASTFSDATKRTKRPSSRQPGANGSAGSLRPOSRG
LVPK"
BASE COUNT      1111 a      866 c      884 g      900 t
ORIGIN

```

```

Query Match      1.5%; Score 148; DB 5; Length 3761;
Best Local Similarity 51.7%; Pred. No. 1.8e-18;
Matches 483; Conservative 0; Mismatches 415; Indels 36; Gaps 5;

```

```

QY 260 gatggagacaagcttcaattcogtgcgtatattcaattcacaagaatcaa 319
DB 367 GAGGTGTCCTCAAAACATTCATTTGTACGACTTATGACTCAACCTCAAGCAAGTGAG 426
QY 320 attaccaagaataagcgtaccatcctatcagatcagcttgcaggatataatggcaca 379
DB 427 CTYATGTGTAAGAACTTTCAGGCCACTGGTGAATCTGTTTGTGCTTGAGACGGGACC 486
QY 380 atattgcatcagcagacacatctcagaagaagcgtacacaaatgatggagaca----- 433
DB 487 ATATTGCTACAGGTCACACAGAACAGAAAGAACTACACATGAGAGAGTGCTGTGT 546
QY 434 ---ccaatlcatggcataataaccaccaagccatacaggagagttttaaatatcag 490
DB 547 GATCCAGAAAAGAGAGGGGTATATACCACTATTGAGCACATATTATCTCACATCTCA 606
QY 491 gagataccgaagaagagttcttctaagagtttctatagagatttacaatgaact 550
DB 607 AGGTCTCAGAACCCAGCACTGCTGTCAGAGCTTCTTACCTAGAGATTATACAGAGAGNA 666
QY 551 gfgaaagacctactgtgatgacagaagaagcccttgaatttcgcgagagattt 610
DB 667 ATCAGAGACTTTTATCTAAGATCAGAG---CAAAAGCCTGAGCTGAAAGAGAGACCT 723
QY 611 aatagaacgltatgttgcgtgacgtgacagaactgttaatgttcttcgaacatgta 670
DB 724 GATATCGGGGTGTATGTCAAGGATTTGCTTCTTTGTTACCAAGAGGTGTGAAGAGATT 783
QY 671 atacagtgatcaaaaaagggtgtaaaaaacagacattatgagagactaaatgaatgat 730
DB 784 GAACATGTGATGAATGTAGGAAATCAAAACCATCTGTTGGAGCGAACCAATATGAATGA 843
QY 731 catagtagctgcatacataatattagaaatgttltgaaagccgagacagaatgat 790
DB 844 CACAGGTCTCGCTCACACAGCCATCTTATGATCACAAATGAGTAGTCAGATTGTTTA 903
QY 791 cccacaatcagaagaactgtatgagagctgtcaatgtatctcacttgaattgttagat 850
DB 904 GATGCGAAGAACCCACATCCGTGTGTA-----AAGCTGAACCTTGTAAGC 948
QY 851 ctgtcgtgagagtgaagaagcagaactgagctgaaagtgtgagacttaagaagaagc 910
DB 949 TTGGCAGGAGGTGAAGGCAAAAGACTGGAGCCCAAGGAGAAAGGCTGAAGAGAGCA 1008
QY 911 tgaacatcaaccgacgctgttattccttggacaggttattaagaagcttagcgagcgc 970
DB 1009 ACTAAGATTAAATCTCTCTCTCCGACATGAGAAATGATCTCTGCTTGGTTGATGG- 1067
QY 971 caggctgtgtgatttataactacagacagaaactcacagaatctccaataat 1030
DB 1068 --GAGAAGCACTCACATCCATACAGGATTCACAACTCACAGGTTACTACAGGACTCA 1125

```

```

QY 1031 ttggaggaatgctaaacgltataatttgacaaattacgccaagttc-----tttt 1084
DB 1126 CTGGGTGGAATGCCAAGACTGTCAATGTTGCCAATATAGGCCAGCTCTCAATGTG 1185
QY 1085 gatgagacttcaagtaacttcagtttgcagtaactgcgcaaatgtggaatactccc 1144
DB 1186 GAGGAGACATTTGACAACTCTCCGCTACTCTCAACAGAGCCCAAAATATCAAGAACCA 1245
QY 1145 catgtaatgaggtccctgagatgataagagcgttgc 1178
DB 1246 AGACTGAATTAAGACCCCAAGATGCTTTATTGC 1279

```

Search completed: October 7, 2002, 14:08:34
Job time: 18586 sec

Tue Oct 8 10:22:05 2002

us-09-150-867-2.rge

Page 21

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 7, 2002, 09:33:18 ; Search time 158 Seconds
(without alignments)
14964.988 Million cell updates/sec

Title: US-09-150-867-2

Perfect score: 9626
Sequence: 1 gaatccgcgagtcgcatag.....ttaaaaaaaaaaacgaattc 9626

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 38353 seqs, 122816752 residues

Word size : 0

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/prodata/2/lna/5A_COMB.seq:*
- 2: /cgn2_6/prodata/2/lna/5B_COMB.seq:*
- 3: /cgn2_6/prodata/2/lna/6A_COMB.seq:*
- 4: /cgn2_6/prodata/2/lna/6B_COMB.seq:*
- 5: /cgn2_6/prodata/2/lna/PCUS_COMB.seq:*
- 6: /cgn2_6/prodata/2/lna/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	0.2	10564	1 US-08-206-176-5	Sequence 5, Appl1
2	19	0.2	423	3 US-08-906-769-134	Sequence 134, App
3	19	0.2	423	3 US-08-906-616-134	Sequence 134, App
4	19	0.2	423	3 US-08-639-075A-134	Sequence 134, App
5	19	0.2	423	4 US-09-012-431-134	Sequence 134, App
6	19	0.2	423	4 US-09-012-692-134	Sequence 134, App
7	19	0.2	423	4 US-08-906-613-134	Sequence 134, App
8	19	0.2	445	1 US-08-207-904-14	Sequence 14, Appl
9	19	0.2	658	4 US-09-328-111-665	Sequence 665, App
10	19	0.2	4080	1 US-08-375-300-1	Sequence 1, Appl1
11	19	0.2	4080	3 US-09-177-431-1	Sequence 1, Appl1
12	19	0.2	4080	5 PCT-US95-16930-1	Sequence 1, Appl1
13	19	0.2	31	2 US-08-859-998-170	Sequence 170, App
14	15	0.2	31	4 US-09-225-928-170	Sequence 170, App
15	18	0.2	315	2 US-08-743-200-1	Sequence 1, Appl1
16	18	0.2	432	4 US-09-001-944-11	Sequence 11, Appl
17	18	0.2	432	4 US-09-240-004A-11	Sequence 11, Appl
18	18	0.2	435	2 US-09-001-944-9	Sequence 9, Appl1
19	18	0.2	435	4 US-09-240-004A-9	Sequence 9, Appl1
20	18	0.2	450	4 US-09-240-004A-13	Sequence 13, Appl1
21	18	0.2	498	2 US-09-001-944-1	Sequence 1, Appl1
22	18	0.2	498	2 US-09-001-944-3	Sequence 3, Appl1
23	18	0.2	498	2 US-09-001-944-5	Sequence 5, Appl1
24	18	0.2	498	2 US-09-001-944-7	Sequence 7, Appl1
25	18	0.2	498	4 US-09-240-004A-1	Sequence 1, Appl1
26	18	0.2	498	4 US-09-240-004A-3	Sequence 3, Appl1
27	18	0.2	498	4 US-09-240-004A-5	Sequence 5, Appl1

28	18	0.2	498	4 US-09-240-004A-7	Sequence 7, Appl1
29	18	0.2	906	4 US-08-976-259-98	Sequence 98, Appl1
30	18	0.2	1558	1 US-08-455-550-7	Sequence 7, Appl1
31	18	0.2	1789	2 US-08-366-490-1	Sequence 1, Appl1
32	18	0.2	1789	3 US-08-860-483A-1	Sequence 1, Appl1
33	18	0.2	1923	1 US-08-077-939-14	Sequence 14, Appl
34	18	0.2	1923	1 US-08-461-599-14	Sequence 14, Appl
35	18	0.2	1923	1 US-08-461-621-14	Sequence 14, Appl
36	18	0.2	1923	1 US-08-465-334-14	Sequence 14, Appl
37	18	0.2	1926	1 US-07-901-703-12	Sequence 12, Appl
38	18	0.2	1926	1 US-08-147-023-26	Sequence 26, Appl
39	18	0.2	1926	1 US-08-278-729A-22	Sequence 22, Appl
40	18	0.2	1926	1 US-08-480-528A-9	Sequence 9, Appl1
41	18	0.2	1926	1 US-08-479-666-9	Sequence 9, Appl1
42	18	0.2	1926	1 US-08-155-343A-22	Sequence 22, Appl
43	18	0.2	1926	1 US-08-406-672-22	Sequence 22, Appl
44	18	0.2	1926	1 US-08-643-563A-22	Sequence 22, Appl
45	18	0.2	1926	1 US-08-447-570-26	Sequence 26, Appl

ALIGNMENTS

RESULT 1
US-08-206-176-5
; Sequence 5, Application US/08206176
; Patent No. 5639940
; GENERAL INFORMATION:
; APPLICANT: Garner, Ian
; APPLICANT: Dalrymple, Michael A
; APPLICANT: Prunkard, Donna E
; TITLE OF INVENTION: Production of fibrinogen in Transgenic
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zymogenetics, Inc.
; STREET: 4225 Roosevelt Way, N.E.
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/206,176
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, Gary E
; REGISTRATION NUMBER: 31-648
; REFERENCE/DOCKET NUMBER: 93-15
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-547-8080 ext 322
; TELEFAX: 206-548-2329
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10564 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: human fibrinogen gamma chain
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(1799..1876, 1973..2017, 2207..2390, 2510
; ..2603, 4211..4341, 4645..4778, 5758..5942, 7426
; ..7703, 9342..9571)
; US-08-206-176-5

Query Match 0.2%; Score 21; DB 1; Length 10564;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2975 TTCTCAAGCAGCTTGAGAA 2995
|||||
DB 3246 TTCTCAAGCAGCTTGAGAA 3266

RESULT 2

US-08-906-769-134
; Sequence 134, Application US/08906769
; Patent No. 6077687
; GENERAL INFORMATION:
; APPLICANT: Grieve, Robert B.
; APPLICANT: Rushlow, Keith E.
; APPLICANT: Wu Hunter, Shirley
; APPLICANT: Frank, Glenn R.
; APPLICANT: Stiegler, Gary
; APPLICANT: Gaines, Patrick J.
; APPLICANT: Silver, Gary
; TITLE OF INVENTION: FLUA PROTEASE PROTEINS, NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES AND USES THEREOF
; NUMBER OF SEQUENCES: 190
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/906,769
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 08/639,075
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-25-C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 134:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 423 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..423
US-08-906-769-134

Query Match 0.2%; Score 19; DB 3; Length 423;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1824 AACTACAGTCAAGAGGA 1842
|||||
DB 222 AACTACAGTCAAGAGGA 240

RESULT 3

US-08-906-616-134
; Sequence 134, Application US/08906616
; Patent No. 6121035
; GENERAL INFORMATION:
; APPLICANT: Grieve, Robert B.
; APPLICANT: Rushlow, Keith E.
; APPLICANT: Wu Hunter, Shirley
; APPLICANT: Frank, Glenn R.
; APPLICANT: Stiegler, Gary
; APPLICANT: Gaines, Patrick J.
; APPLICANT: Silver, Gary
; TITLE OF INVENTION: FLUA AMINOPEPTIDASE PROTEINS AND USES THEREOF
; NUMBER OF SEQUENCES: 190
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Sheridan Ross P.C.
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/906,616
; FILING DATE: 05-AUG-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2618-25-C2-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 134:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 423 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..423
US-08-906-616-134

Query Match 0.2%; Score 19; DB 3; Length 423;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1824 AACTACAGTCAAGAGGA 1842
|||||
DB 222 AACTACAGTCAAGAGGA 240

RESULT 4

US-08-639-075A-134
; Sequence 134, Application US/08639075A
; Patent No. 6150125
; GENERAL INFORMATION:
; APPLICANT: Grieve, Robert B.
; APPLICANT: Rushlow, Keith E.
; APPLICANT: Wu Hunter, Shirley
; APPLICANT: Frank, Glenn R.
; APPLICANT: Stiegler, Gary
; APPLICANT: Gaines, Patrick J.
; APPLICANT: Silver, Gary
; TITLE OF INVENTION: FLUA PROTEASE PROTEINS, NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES AND USES THEREOF
; NUMBER OF SEQUENCES: 190
; CORRESPONDENCE ADDRESSES:
;

ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/639,075A
FILING DATE: 24-APR-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 134:
SEQUENCE CHARACTERISTICS:
LENGTH: 423 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..423
US-08-639-075A-134

Query Match 0.2%; Score 19; DB 3; Length 423;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1824 AACTACAGTCAAGAGAGA 1842
DB 222 AACTACAGTCAAGAGAGA 240

RESULT 5
US-09-012-431-134
Sequence 134, Application US/09012431
Patent No. 6180383
GENERAL INFORMATION:
APPLICANT: Grieve, Robert B.
Rushlow, Keith E.
Wu Hunter, Shirley
Frank, Glenn R.
Stiegler, Gary
Gaines, Patrick J.
Silver, Gary
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
MOLECULES AND USES THEREOF
NUMBER OF SEQUENCES: 190
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/012,431
FILING DATE: 23-Jan-1998

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/639,075
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 134:
SEQUENCE CHARACTERISTICS:
LENGTH: 423 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..423
US-09-012-431-134

Query Match 0.2%; Score 19; DB 4; Length 423;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1824 AACTACAGTCAAGAGAGA 1842
DB 222 AACTACAGTCAAGAGAGA 240

RESULT 6
US-09-012-692-134
Sequence 134, Application US/09012692
Patent No. 6214579
GENERAL INFORMATION:
APPLICANT: Grieve, Robert B.
Rushlow, Keith E.
Wu Hunter, Shirley
Frank, Glenn R.
Stiegler, Gary
Gaines, Patrick J.
Silver, Gary
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
MOLECULES AND USES THEREOF
NUMBER OF SEQUENCES: 190
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/012,692
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/639,075
FILING DATE: 24-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700

TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 134:
SEQUENCE CHARACTERISTICS:
LENGTH: 423 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..423
US-09-012-692-134

Query Match 0.2%; Score 19; DB 4; Length 423;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1824 AACTACAGTCAAGAGGA 1842
|||||
DB 222 AACTACAGTCAAGAGGA 240

RESULT 7
US-08-906-613-134
Sequence 134, Application US/08906613
Patent No. 6232096
GENERAL INFORMATION:
APPLICANT: Grieve, Robert B.
APPLICANT: Rushlow, Keith E.
APPLICANT: Wu Hunter, Shirley
APPLICANT: Frank, Glenn R.
APPLICANT: Stiegler, Gary
APPLICANT: Gaines, Patrick J.
APPLICANT: Silver, Gary
TITLE OF INVENTION: FIEN PROTEASE PROTEINS, NUCLEIC ACID
TITLE OF INVENTION: MOLECULES AND USES THEREOF
NUMBER OF SEQUENCES: 190
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/906.613
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/639,075
FILING DATE: 24-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 134:
SEQUENCE CHARACTERISTICS:
LENGTH: 423 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..423

US-08-906-613-134

Query Match 0.2%; Score 19; DB 4; Length 423;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1824 AACTACAGTCAAGAGGA 1842
|||||
DB 222 AACTACAGTCAAGAGGA 240

RESULT 8
US-08-207-904-14
Sequence 14, Application US/08207904
Patent No. 5477002
GENERAL INFORMATION:
APPLICANT: Tuttle, AnnMarie
APPLICANT: Crossland, Lyle D.
TITLE OF INVENTION: Antier-Specific cDNA Sequences, Genomic
TITLE OF INVENTION: DNA Sequences and Recombinant DNA Sequences
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/207,904
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/908,242
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Lazar, Steven R.
REGISTRATION NUMBER: 32,618
REFERENCE/DOCKET NUMBER: CGC 1624
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8615
TELEFAX: (919)541-8615
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 445 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Nicotiana tabacum
INDIVIDUAL ISOLATE: Ant68
FEATURE:
NAME/KEY: CDS
LOCATION: 2..445
US-08-207-904-14

Query Match 0.2%; Score 19; DB 1; Length 445;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4924 CAATGTCAATACGCA 4942
|||||
DB 301 CAATGTCAATACGCA 319

RESULT 9
US-09-328-111-665
Sequence 665, Application US/09328111
Patent No. 6262333
GENERAL INFORMATION:
APPLICANT: Endege, Wilson O.
APPLICANT: Steinmann, Kathleen E.
APPLICANT: Astle, Jon H.
APPLICANT: Burgess, Christopher C.
APPLICANT: Bushnell, Steven E.
APPLICANT: Carroll III, Eddie
APPLICANT: Catino, Theodore J.
APPLICANT: Dertli, Adnan
APPLICANT: Ford, Donna M.
APPLICANT: Lewis, Marcia E.
APPLICANT: Monahan, John E.
APPLICANT: Schlegel, Robert
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
FILE REFERENCE: CCD-257 (US)
CURRENT APPLICATION NUMBER: US/09/328,111
CURRENT FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: US 60/088,801
EARLIER FILING DATE: 1998-06-10
NUMBER OF SEQ ID NOS: 850
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 665
LENGTH: 658
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(658)
OTHER INFORMATION: n = A,T,C or G
US-09-328-111-665

Query Match 0.2%; Score 19; DB 4; Length 658;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3813 TGCAGCTGAGAACTCTGA 3831
|||||
Db 24 TGCAGCTGAGAACTCTGA 42

RESULT 10
US-08-375-300-1
Sequence 1, Application US/08375300
Patent No. 5679566
GENERAL INFORMATION:
APPLICANT: Feng, He
APPLICANT: Jacobson, Allan S.
TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street Suite 3100
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/375,300
FILING DATE: 20-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Fasse, J. P.
REGISTRATION NUMBER: 32,983

REFERENCE/DOCKET NUMBER: 04020/046001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)542-5070
TELEFAX: (617)542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4080 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-375-300-1

Query Match 0.2%; Score 19; DB 1; Length 4080;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4166 GAAAGAGTGAAGCTTCAA 4184
|||||
Db 862 GAAAGAGTGAAGCTTCAA 880

RESULT 11
US-09-177-431-1
Sequence 1, Application US/09177431
Patent No. 6071700
GENERAL INFORMATION:
APPLICANT: He, Feng
APPLICANT: Jacobson, Allan S.
TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE PRODUCTION IN THE
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/177,431
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/955,472
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Fasse, J. Peter
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 07917/050001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-9806
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4080 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-177-431-1

Query Match 0.2%; Score 19; DB 3; Length 4080;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4166 GAAAGAGTGAAGCTTCAA 4184

Db 862 GAAAGAGATGAGCTCAA 880

RESULT 12
PCT-US95-16930-1

Sequence 1, Application PC/TUS9516930
GENERAL INFORMATION:
APPLICANT: UNIVERSITY OF MASSACHUSETTS MEDICAL SCHOOL
TITLE OF INVENTION: HETEROLOGOUS POLYPEPTIDE
TITLE OF INVENTION: PRODUCTION IN THE ABSENCE OF
TITLE OF INVENTION: NONSENSE-MEDIATED mRNA DECAY
TITLE OF INVENTION: FUNCTION
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street Suite 3100
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/16930
FILING DATE: 27-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/375,300
FILING DATE: 20-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Fasse, J. Peter
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 04020/046W01
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)542-5070
TELEFAX: (617)542-8906
TELEX: 200154
INFORMATION FOR SEQ. ID NO.: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4080 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
PCT-US95-16930-1

Query Match Best Local Similarity 0.2%; Score 19; DB 5; Length 4080;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4166 GAAAGAGATGAGCTCAA 4184
Db 862 GAAAGAGATGAGCTCAA 880

RESULT 13
US-08-859-998-170

Sequence 170, Application US/08859998
Patent No. 5994076
GENERAL INFORMATION:
APPLICANT: Chenchik, Alex
APPLICANT: Bibilashvili, Robert
TITLE OF INVENTION: METHOD OF ASSAYING DIFFERENTIAL
TITLE OF INVENTION: EXPRESSION
NUMBER OF SEQUENCES: 1375
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 2200 Sand Hill Road, Suite 100
CITY: Menlo Park

STATE: CA
COUNTRY: US
ZIP: 94025

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/859,998
FILING DATE: 21-MAY-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Field, Bret E.
REGISTRATION NUMBER: 37,620
REFERENCE/DOCKET NUMBER: 09096/002001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-322-5070
TELEFAX: 415-854-0875
INFORMATION FOR SEQ. ID NO.: 170:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
OTHER INFORMATION: oligonucleotide primer
US-08-859-998-170

Query Match Best Local Similarity 0.2%; Score 18; DB 2; Length 31;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5338 ATGTCTGCTAATGAGAA 5355
Db 7 ATGTCTGCTAATGAGAA 24

RESULT 14
US-09-225-928-170

Sequence 170, Application US/09225928
Patent No. 6352829
GENERAL INFORMATION:
APPLICANT: Chenchik, Alex
APPLICANT: Bibilashvili, Robert
TITLE OF INVENTION: METHOD OF ASSAYING DIFFERENTIAL
TITLE OF INVENTION: EXPRESSION
NUMBER OF SEQUENCES: 1375
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson, P.C.
STREET: 2200 Sand Hill Road, Suite 100
CITY: Menlo Park
STATE: CA
COUNTRY: US
ZIP: 94025
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/225,928
FILING DATE: 05-Jan-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/859,998
FILING DATE: 21-MAY-1997
ATTORNEY/AGENT INFORMATION:

NAME: Field, Bret E.
REGISTRATION NUMBER: 37,620
REFERENCE/DOCKET NUMBER: 09096/002001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-322-5070
TELEFAX: 415-854-0875
INFORMATION FOR SEQ ID NO: 170:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
OTHER INFORMATION: oligonucleotide primer
SEQUENCE DESCRIPTION: SEQ ID NO: 170:
US-09-225-928-170

Query Match 0.2%; Score 18; DB 4; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5338 ATGCTGCTAATGAGAA 5355
|||||
Db 7 ATGCTGCTAATGAGAA 24

RESULT 15
US-08-743-200-1
Sequence 1, Application US/08743200
Patent No. 5861260
GENERAL INFORMATION:
APPLICANT: Doherty, Stephen J.
TITLE OF INVENTION: DIAGNOSTIC METHODS FOR SCREENING
TITLE OF INVENTION: PATIENTS FOR SCLERODERMA
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/743,200
FILING DATE: 05-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Fasse, J. Peter
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 07917/025001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 315 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 1...315
US-08-743-200-1

Query Match 0.2%; Score 18; DB 2; Length 315;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4408 ATTCAGTTAGAGAGGT 4425
|||||
Db 24 ATTCAGTTAGAGAGGT 41

Search completed: October 7, 2002, 11:14:00
Job time : 235 secs

THIS PAGE BLANK (USPTO)

FEATURES

Location/Qualifiers

1. 554

/organism="Xenopus laevis"

/db_xref="taxon:8355"

/clone="PBX0105E04"

/clone_lib="Blackshear/Soares normalized Xenopus egg library"

/sex="female"

/tissue_type="unfertilized egg"

/cell_type="unfertilized egg"

/dev_stage="unfertilized egg"

/lab_host="DH10B"

/note="Vector: pT73-Pac; Site 1: EcoRI; Site 2: NotI; PolyA-selected mRNA was prepared from unfertilized Xenopus laevis eggs. The library was constructed in the vector pT73-Pac as described in Bonaldo, M.F., Lennon, G. and Soares, M.B. 'Normalization and subtraction: two approaches to facilitate gene discovery', Genome Research 6:791-806, 1996. The first strand synthesis used a NotI-dt18 primer; double stranded cDNAs were ligated to EcoRI adapters, digested with NotI, and directionally cloned into the NotI and EcoRI-digested pT73-Pac vector. The library contained approximately 7.2 x 10⁵ recombinants, with average insert sizes of 1-1.5 kb."

BASE COUNT 177 a 107 c 129 g 141 t

ORIGIN

Query Match 5.6%; Score 541; DB 9; Length 554;
Best Local Similarity 100.0%; Pred. No. 2.2e-241;
Matches 541; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6662 GCTGTGATTCAGAGCAACACATTTGCTCAGAGCTCTCTAGAGAGCTTAAAGGAA 7021
DB 14 GCTGTGATTCAGAGCAACACATTTGCTCAGAGCTCTCTAGAGAGCTTAAAGGAA 73
QY 7022 ACTGAAGCACAAACATTCATGTTAAATPAAAGAACTCTCTCATCCACGCTCC 7081
DB 74 ACTGAAGCACAAACATTCATGTTAAATPAAAGAACTCTCTCATCCACGCTCC 133
QY 7082 AGATCTTTGGGACCTTGCAGCACTGATGTAAGCTAAATACCTCAGACCTT 7141
DB 134 AGATCTTTGGGACCTTGCAGCACTGATGTAAGCTAAATACCTCAGACCTT 193
QY 7142 CTGAACAATTTAAGGTTGTATACGAACTGCTGAGTCAAGAAATCATAGCTTATC 7201
DB 194 CTGAACAATTTAAGGTTGTATACGAACTGCTGAGTCAAGAAATCATAGCTTATC 253
QY 7202 AAAGATTATGAGAGAGACCTTGTCTGAGCAAAAAGGAGTATGAGCTGCATCCAA 7261
DB 254 AAAGATTATGAGAGAGACCTTGTCTGAGCAAAAAGGAGTATGAGCTGCATCCAA 313
QY 7262 CTGAGATTTTGGGACGACGCAAGAAATGTCGGATTCTGCATCTGAGAACTCAAG 7321
DB 314 CTGAGATTTTGGGACGACGCAAGAAATGTCGGATTCTGCATCTGAGAACTCAAG 373
QY 7322 TTCTGTGAATTTGAATCTTGAATGATCTTTTAAAGCAAAATATATTCAGAGT 7381
DB 374 TTCTGTGAATTTGAATCTTGAATGATCTTTTAAAGCAAAATATATTCAGAGT 433
QY 7382 GTCAGAGATGATCTTTAGAGGTGACAGTATTCCTAAATCAGTAGATCAACACGCAA 7441
DB 434 GTCAGAGATGATCTTTAGAGGTGACAGTATTCCTAAATCAGTAGATCAACACGCAA 493
QY 7442 GAAGAGCTTGAGCAAGAAAGCTTTATGAGAGTGTGGAGAAATTTGGAGATCTGCAC 7501
DB 494 GAAGAGCTTGAGCAAGAAAGCTTTATGAGAGTGTGGAGAAATTTGGAGATCTGCAC 553
QY 7502 G 7502
DB 554 G 554

RESULT 2

BI349745

LOCUS

DEFINITION

BI349745 426 bp mRNA linear EST 31-JUL-2001
daee3c09.y3 Blackshear/Soares normalized Xenopus egg library
Xenopus laevis cDNA clone IMAGE:4678672 5' similar to TR:042263
O42263 KINESIN-RELATED PROTEIN. ; mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

Clifton, S., Johnson, S.L., Blumberg, B., Song, J., Hillier, L., Pape, D., Martin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y., Person, B., Gibbons, M., Harvey, N., Rittler, E., Jackson, Y., McCann, R., Waterston, R. and Wilson, R.
Washu Xenopus EST project, 1999
Unpublished (1999)
Contact: Sandy Clifton, Ph.D.
Washu Xenopus EST project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

TITLE

JOURNAL

COMMENT

Email: est@watson.wustl.edu
Library constructed by Bento Soares and M. Fatima Bonaldo (University of Iowa). DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: Xenopus clones from this library are available through the I.M.A.G.E. Consortium/LINL at: info@image.llnl.gov
Putative full length read
vector to vector length is 496.

FEATURES

source

Location/Qualifiers

1. 426

/organism="Xenopus laevis"

/db_xref="taxon:8355"

/clone="IMAGE:4678672"

/clone_lib="Blackshear/Soares normalized Xenopus egg library"

/sex="female"

/tissue_type="unfertilized egg"

/cell_type="unfertilized egg"

/dev_stage="unfertilized egg"

/lab_host="DH10B"

/note="Vector: pT73-Pac; Site 1: EcoRI; Site 2: NotI; PolyA-selected mRNA was prepared from unfertilized Xenopus laevis eggs. The library was constructed in the vector pT73-Pac as described in Bonaldo, M.F., Lennon, G. and Soares, M.B. 'Normalization and subtraction: two approaches to facilitate gene discovery', Genome Research 6:791-806, 1996. The first strand synthesis used a NotI-dt18 primer; double stranded cDNAs were ligated to EcoRI adapters, digested with NotI, and directionally cloned into the NotI and EcoRI-digested pT73-Pac vector. The library contained approximately 7.2 x 10⁵ recombinants, with average insert sizes of 1-1.5 kb."

BASE COUNT 145 a 73 c 105 g 103 t

ORIGIN

Query Match 4.4%; Score 424; DB 10; Length 426;
Best Local Similarity 100.0%; Pred. No. 1.2e-186;
Matches 424; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7239 GGCATGATGAGTGCAGCTCCAAAGCTGTTGGAGCAGCAGCGCAAGAAATGTCGG 7298
DB 1 GGCATGATGAGTGCAGCTCCAAAGCTGTTGGAGCAGCAGCGCAAGAAATGTCGG 60
QY 7299 ATTCGATCTGAGAGAACTCAAGTTCTGTGAATTTGAATCTTGAATGATGTTACTTTTA 7358
DB 61 ATTCGATCTGAGAGAACTCAAGTTCTGTGAATTTGAATGATGTTACTTTTA 120


```

OY 7359 AAAAGCAATATATATTCAGAGTGTCCAGATGACTTTTCAGAGTGTGACGATATCTCAA 7418
DB 121 AAAAGCAATATATATTCAGAGTGTCCAGATGACTTTTCAGAGTGTGACGATATCTCAA 180
OY 7419 ATCAAGTAGCATCAACACTGCAAGAGAGCTTGAGCACAAGAAAGGCTTTATGCACTGCT 7478
DB 181 ATCAAGTAGCATCAACACTGCAAGAGAGCTTGAGCACAAGAAAGGCTTTATGCACTGCT 240
OY 7479 TGGAGGAATTTGGAGATCTGCACGCTGCATGCTAAGAAACTGATGAAGGATGCAACAGG 7538
DB 241 TGGAGGAATTTGGAGATCTGCACGCTGCATGCTAAGAAACTGATGAAGGATGCAACAGG 300
OY 7539 AAAATAGCGCATGCTCTTACCATATACAGCTCTTAACAAAAGGCTAAAGGCACTGTTTC 7598
DB 301 AAAATAGCGCATGCTCTTACCATATACAGCTCTTAACAAAAGGCTAAAGGCACTGTTTC 360
OY 7599 AGTCGAAAATACACGTGATACCGTATATCTGCAACGAGTTTGAAGGAAATTTGCAAG 7658
DB 361 AGTCGAAAATACACGTGATACCGTATATCTGCAACGAGTTTGAAGGAAATTTGCAAG 420
OY 7659 AGAA 7662
DB 421 AGAA 424

RESULT 3
BM262378 636 bp mRNA linear EST 18-DEC-2001
LOCUS daq40c12.y3 Blackshear/Soares normalized Xenopus egg library
DEFINITION Xenopus laevis cDNA clone IMAGE:4783871 5' similar to TR:042263
O42263 KINESIN-RELATED PROTEIN.; mRNA sequence.
ACCESSION BM262378
VERSION BM262378.1 GI:17925418
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 636)
AUTHORS Clifton,S., Johnson,S.L., Blumberg,B., Song,J., Hillier,L., Pape,D.,
Martin,J., Wylie,T., Underwood,K., Theising,B., Bowers,Y., Person
B., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R.,
Waterston,R. and Wilson,R.
Washu Xenopus EST project, 1999
Unpublished (1999)
Other_ESTS: daq40c12.x3
Contact: Sandy Clifton, Ph.D.
Washu Xenopus EST project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Library constructed by Bento Soares and M. Fatima Bonaldo
(University of Iowa). DNA Sequencing by: Washington University
Genome Sequencing Center
Clone distribution: Xenopus clones from this library are available
through the I.M.A.G.E. Consortium/BLML at: info@image.llnl.gov
Seq primer: -40RP from Glibco
High quality sequence stop: 493.
FEATURES
Source
Location/Qualifiers
1..636
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="IMAGE:4783871"
/library="
lib="Blackshear/Soares normalized Xenopus egg
library"
/sex="female"
/tissue_type="unfertilized egg"
/cell_type="unfertilized egg"
/dev_stage="unfertilized egg"
/lab_host="DH10B"

```

```

/note="Vector: pT773-Pac; Site.1: EcoRI; Site.2: NotI;
PolyA-selected mRNA was prepared from unfertilized Xenopus
laevis eggs. The library was constructed from unfertilized Xenopus
pT773-Pac as described in Bonaldo, M.F., Lennon, G. and
Soares, M.B. 'Normalization and subtraction: two
approaches to facilitate gene discovery', Genome Research
6:791-806, 1996. The first strand synthesis used a
NotI-dT18 primer: double stranded cDNAs were ligated to
EcoRI adapters, digested with NotI, and directionally
cloned into the NotI and EcoRI-digested pT773-Pac vector.
The library contained approximately 7.2 x 105
recombinants, with average insert sizes of 1-1.5 kb."
BASE COUNT 185 a 126 c 125 g 200 t
ORIGIN
Query Match 4.3%; Score 417; DB 10; Length 636;
Best Local Similarity 100.0%; Pred. No. 2.2e-183;
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 8876 CAGTTTGGATATATCTAAGCTTGGTGAATTTTCAGAGCTCAACACGAGACCAAT 8935
DB 97 CAGTTTGGATATATCTAAGCTTGGTGAATTTTCAGAGCTCAACACGAGACCAAT 156
OY 8936 GACAAAGTCAGGCTGAGAACTGCTGATGAAGCAAAAAAGAAAGGCACTGATGT 8995
DB 157 GACAAAGTCAGGCTGAGAACTGCTGATGAAGCAAAAAAGAAAGGCACTGATGT 216
OY 8996 AAAACATCTAGATCCCTGATACATCTGCTCTCTGCAAAAGACTGCTACTCT 9055
DB 217 AAAACATCTAGATCCCTGATACATCTGCTCTCTGCAAAAGACTGCTACTCT 276
OY 9056 GCCTTTGTAGGAAGAAACACACAGAACTGCTGCTGCAAAAGACTGCTACTGCA 9115
DB 277 GCCTTTGTAGGAAGAAACACACAGAACTGCTGCTGCAAAAGACTGCTACTGCA 336
OY 9116 AGCAAAAGTTGTTCTTAGTAAATACATGCTGCTGCGAGAGTGTACGCTTAAATA 9175
DB 337 AGCAAAAGTTGTTCTTAGTAAATACATGCTGCTGCGAGAGTGTACGCTTAAATA 396
OY 9176 AAGTGAATAGCTGACGCTGATATATGTTATGTTATGTTATGTTATGTTATGTTAT 9235
DB 397 AAGTGAATAGCTGACGCTGATATATGTTATGTTATGTTATGTTATGTTATGTTAT 456
OY 9236 AGTCCCTTACATCTATTTGGCTCATCTGATTTCTTTGTTGGCCACACACA 9292
DB 457 AGTCCCTTACATCTATTTGGCTCATCTGATTTCTTTGTTGGCCACACACA 513

RESULT 4
BM262071/c 583 bp mRNA linear EST 18-DEC-2001
LOCUS daq40c12.x3 Blackshear/Soares normalized Xenopus egg library
DEFINITION Xenopus laevis cDNA clone IMAGE:4783871 3', mRNA sequence.
ACCESSION BM262071
VERSION BM262071.1 GI:17925111
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 583)
AUTHORS Clifton,S., Johnson,S.L., Blumberg,B., Song,J., Hillier,L., Pape,D.,
Martin,J., Wylie,T., Underwood,K., Theising,B., Bowers,Y., Person
B., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R.,
Waterston,R. and Wilson,R.
Washu Xenopus EST project, 1999
Unpublished (1999)
Contact: Sandy Clifton, Ph.D.
Washu Xenopus EST project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800

```

Fax: 314 286 1810
 Email: estewatson.wustl.edu
 Library constructed by Bento Soares and M. Farina Bonaldo
 (University of Iowa). DNA Sequencing by: Washington University
 Genome Sequencing Center
 Clone distribution: Xenopus clones from this library are available
 through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
 Seq primer: -40UP from GIBCO
 High quality sequence stop: 510.

FEATURES
 source
 1. 583
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /clone="IMAGE:4783871"
 /clone_lib="Blackshear/Soares normalized Xenopus egg
 library"
 /sex="female"
 /tissue_type="unfertilized egg"
 /cell_type="unfertilized egg"
 /dev_stage="unfertilized egg"
 /lab_host="DH10B"
 /note="Vector: p773-Pac; Site_1: EcoRI; Site_2: NotI;
 Polya-selected mRNA was prepared from unfertilized Xenopus
 laevis eggs. The library was constructed in the vector
 p773-Pac as described in Bonaldo, M.F., Lennon, G. and
 Soares, M.B. 'Normalization and subtraction: two
 approaches to facilitate gene discovery', Genome Research
 6:791-806, 1996. The first strand synthesis used a
 NotI-dn18 primer: double stranded cDNAs were ligated to
 EcoRI adapters, digested with NotI, and directionally
 cloned into the NotI and EcoRI-digested p773-Pac vector.
 The library contained approximately 7.2 x 10⁵
 recombinants, with average insert sizes of 1-1.5 kb."

BASE COUNT
 201 a 103 c 99 g 180 t

ORIGIN

Query Match
 Best Local Similarity 100.0%; Score 228; DB 10; Length 583;
 Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 9065 TAGGAAGAACACTAGAAAGTGCATTAAGAGAGCTCTGCGAAGCAAAAGT 9124
 |||||||
 Db 583 TAGGAAGAACACTAGAAAGTGCATTAAGAGAGCTCTGCGAAGCAAAAGT 9124
 Qy 9125 TGTCTTCTAGTAATCACTGCTGGCGAGTGTACGCTTTAAATAAAGTCAAT 9184
 |||||||
 Db 523 TGTCTTCTAGTAATCACTGCTGGCGAGTGTACGCTTTAAATAAAGTCAAT 9184
 Qy 9185 AGCTCCACGCTCTATTATATGTTATGCTGTATTTGTACACTTTAAGTCCCTTG 9244
 |||||||
 Db 463 AGCTCCACGCTCTATTATATGTTATGCTGTATTTGTACACTTTAAGTCCCTTG 9244
 Qy 9245 ACTTCATATTGGCTCATCTGATGTTCTTTGTTGCGACACACA 9292
 |||||||
 Db 403 ACTTCATATTGGCTCATCTGATGTTCTTTGTTGCGACACACA 356

RESULT 5
 BG408399 430 bp mRNA linear EST 13-MAR-2001
 LOCUS
 DEFINITION
 dd01b05.y1 NICHD XGC 001 Xenopus laevis cDNA clone IMAGE:3405249 5'
 sequence.
 accession
 BG408399
 version
 BG408399.1 GI:13314748
 keywords
 EST.
 source
 African clawed frog.
 organism
 Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 Xenopodinae; Xenopus.
 reference
 1 (bases 1 to 430)
 authors
 NCI-GAP <http://www.ncbi.nlm.nih.gov/ncigap>.

TITLE
 JOURNAL
 COMMENT
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: c9aps-remail.nih.gov
 Tissue Procurement: Martha Rebert, Steven L. Klein, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Clone distribution: Xenopus clones from this library are available
 through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
 Seq primer: -40RP from GIBCO
 High quality sequence stop: 216.

FEATURES
 source
 1. 430
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /clone="IMAGE:3405249"
 /clone_lib="NICHD XGC 001"
 /tissue_type="oocytes"
 /lab_host="DH10B (phage-resistant)"
 /note="Vector: PCMV-SPO86; Site_1: NotI; Site_2: SalI;
 Cloned unidirectionally. Primer: Oligo dt. Average insert
 size 2.2 kb. Constructed by Life Technologies."

BASE COUNT
 167 a 76 c 105 g 81 t

ORIGIN

Query Match
 Best Local Similarity 99.6%; Score 181; DB 10; Length 430;
 Matches 231; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 7912 AAACGCAATCAGGTGAACCTAATAAAGCAATATCTCAGCGTCAATGAAGA 7971
 |||||||
 Db 1 AAACGCAATCAGGTGAACCTAATAAAGCAATATCTCAGCGTCAATGAAGA 60
 Qy 7972 AATGAAACCTACGACAAAATGTTGCCAAGTGCATACATATTAAGAAGAAATGA 8031
 |||||||
 Db 61 AATGAAACCTACGACAAAATGTTGCCAAGTGCATACATATTAAGAAGAAATGA 120
 Qy 8032 CAACCTTAAACCTAGAGTGTAAAGATGAATGAATGAATGAATGAATGA 8091
 |||||||
 Db 121 CAACCTTAAACCTAGAGTGTAAAGATGAATGAATGAATGAATGAATGA 180
 Qy 8092 AGACCAAGAGATGCTCTACTTAAGCTTTGTTGCAAGATGAAGAGAGCC 8143
 |||||||
 Db 181 AGACCAAGAGATGCTCTACTTAAGCTTTGTTGCAAGATGAAGAGAGCC 232

RESULT 6
 BG371437/c 545 bp mRNA linear EST 09-MAR-2001
 LOCUS
 DEFINITION
 dac48h12.x1 RIKEN Xenopus egg Xenopus laevis cDNA clone
 IMAGE:4433519 3' similar to TR:042263 042263 KINSIN-RELATED
 PROTEIN. ; mRNA sequence.
 accession
 BG371437
 version
 BG371437.1 GI:13267974
 keywords
 EST.
 source
 African clawed frog.
 organism
 Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 Xenopodinae; Xenopus.
 reference
 1 (bases 1 to 545)
 authors
 Clifton, S., Johnson, S.L., Blumberg, B., Song, J., Hallier, L., Pape, D.,
 Martin, J., Wylie, T., Underwood, K., Theisling, B., Bowers, Y., Person
 Waterston, R. and Wilson, N., Rittler, E., Jackson, Y., McCann, R.,
 Washu Xenopus EST project, 1999
 unpublished (1999)
 contact: Sandy Clifton, Ph.D.
 Washu Xenopus EST project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800
 Fax: 314 286 1810
 Email: estewatson.wustl.edu
 Library constructed by Piero Carninci (RIKEN, Genome Science Laboratory, Japan: <http://genome.riken.go.jp/>) using egg RNA supplied by Marc Kirschner (Harvard Medical School, Dept. of Cell Biology) DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: Xenopus clones from this library are available through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
 Seq primer: -40UP from G1bco
 High quality sequence stop: 376.
 Location/Qualifiers

FEATURES

1..545
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /clone="IMAGE:443519"
 /clone_1lb="RIKEN Xenopus egg"
 /tissue_type="egg (unfertilized)"
 /lab_host="DH10B"
 /note="Vector: PCS2G (custom); Site_1: BamHI; Site_2: XhoI; CDNA Oligo-dT primed using oligo 5'-GAGAGAGAGCATCCATCTCGAGAGCTTTTCTTTTCTTNN-3' and directionally cloned into XhoI and BamHI sites using the 5' adaptor 5'-AGAGAGAGCTCGAGCTCTATAGTGACACTATAGACCA-3'. Size selection was done automatically by the cloning vector (lambda f1C-II, P. Carninci et al., unpublished). Library was amplified at the phage stage and normalized (ROF value = 5). This library was constructed using the Captrapper method and Trehalose thermoactivated reverse transcriptase to enrich for full-length clones. Library constructed by Piero Carninci (RIKEN, Genome Science Laboratory, Japan: <http://genome.riken.go.jp/>) using egg RNA supplied by Marc Kirschner (Harvard Medical School, Dept. of Cell Biology). References: Methods Enzymol. 303:19-44 and Genome Res. 10:1617-1630 Please contact Marc Kirschner (marchms.harvard.edu) for information on obtaining aliquots of this cDNA library."

BASE COUNT 139 a 104 c 108 g 194 t
 ORIGIN

Query Match 1.9%; Score 179; DB 10; Length 545;
 Best Local Similarity 99.1%; Pred. No. 4.6e-72;
 Matches 329; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1508 GCAAGTCTCTGACATGCTTCATTTCCAGAAATTGATGACTCTGTTGTACAGAGTTT 1567
 |||
 DB 338 GCAAGTCTCTGACATGCTTCATTTCCAGAAATTGATGACTCTGTTGTACAGAGTTT 279
 QY 1568 TCTGATTTTGATGAGCCCTCTTCATGATGAGACGCAATGGAATAGATGACAGATGGAAT 1627
 |||
 DB 278 TTTGATTTTGATGAGCCCTCTTCATGATGAGACGCAATGGAATAGATGAGATGGAAT 219
 QY 1628 TTAGCCAGTAAAGTAACTACAGAGAAAAGACATCTTCATCAATCAATGATAGACTTT 1687
 |||
 DB 218 TTAGCCAGTAAAGTAACTACAGAGAAAAGACATCTTCATCAATCAATGATAGACTTT 159
 QY 1688 GGACAGATTTCGACAGTGTTCAGTTTCATGATTTCTTAAGGAAAAACAGCTACATAC 1747
 |||
 DB 158 GGACAGATTTCGACAGTGTTCAGTTTCATGATTTCTTAAGGAAAAACAGCTACATAC 99
 QY 1748 CTCGCCAAGAGACTCTGGTATATGGTGAATGAGAAAAGCTTTTGTAAAAAGAGATC 1807
 |||
 DB 98 CTCGCCAAGAGACTCTGGTATATGGTGAATGAGAAAAGCTTTTGTAAAAAGAGATC 39
 QY 1808 ACAAGCTCCAGCAACACTACAGTCAAGA 1839
 |||
 DB 38 ACAAGCTCCAGCAACACTACAGTCAAGA 7

RESULT 7
 BG486505/c

LOCUS BG486505 408 bp mRNA linear EST 22-MAR-2001
 DEFINITION dd01b05.x1 NICHD XGC 001 Xenopus laevis cDNA clone IMAGE:3405249
 3', mRNA sequence.
 ACCESSION BG486505
 VERSION BG486505.1 GI:13434082
 KEYWORDS EST.
 SOURCE African clawed frog.
 ORGANISM Xenopus laevis

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
 xenopodinae; Xenopus.
 1 (bases 1 to 408)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 Other ESTs: dd01b05.y1
 Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov

JOURNAL

Tissue Procurement: Martha Redbert, Steven L. Klein, Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: Xenopus clones from this library are available through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
 Seq primer: -40UP from G1bco
 High quality sequence stop: 320.
 Location/Qualifiers

FEATURES

1..408
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /clone="IMAGE:3405249"
 /clone_1lb="NICHD XGC 001"
 /tissue_type="oocytes"
 /lab_host="DH10B (phage-resistant)"
 /note="Vector: PCMV-SPORE6; Site_1: NotI; Site_2: SalI;
 cloned unidirectionally. Primer: Oligo dT. Average insert size 2.2 kb. Constructed by Life Technologies."

BASE COUNT 149 a 62 c 70 g 127 t
 ORIGIN

Query Match 1.8%; Score 172; DB 10; Length 408;
 Best Local Similarity 100.0%; Pred. No. 9e-69;
 Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9290 ACACACTGCTGAAAGATGACATTTGCAGTGTATGTTTACTGACTGCTCTGGGCG 9349
 |||
 DB 334 ACACACTGCTGAAAGATGACATTTGCAGTGTATGTTTACTGACTGCTCTGGGCG 275
 QY 9350 CATCATGCAAAACACACATTAGTGTCCATGTTTTCATCTATTATATGCTGCAC 9409
 |||
 DB 274 CATCATGCAAAACACACATTAGTGTCCATGTTTTCATCTATTATATGCTGCAC 215
 QY 9410 TTGTGAAATAGAAATACTACAAAAGATGAGCAAAAGTTGTTTCTCACT 9461
 |||
 DB 214 TTGTGAAATAGAAATACTACAAAAGATGAGCAAAAGTTGTTTCTCACT 163

RESULT 8
 AM636063 457 bp mRNA linear EST 26-APR-2001
 LOCUS b142c09.w1 Blackshear/Soares normalized Xenopus egg library Xenopus
 DEFINITION laevis cDNA clone PBX0042C09 5', mRNA sequence.
 ACCESSION AM636063
 VERSION AM636063.1 GI:7393144
 KEYWORDS EST.
 SOURCE African clawed frog.
 ORGANISM Xenopus laevis

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
 xenopodinae; Xenopus.
 1 (bases 1 to 457)
 Blackshear, P.J., Lal, W.S., Thorn, J.M., Kennington, E.A., Staffa, N.G.

J.F., Moore, D.T., Bouffard, G.G., Beckstrom-Sternberg, S.M., Touchman, J.W., Bonaldo, M.F., and Soares, M.B.
The NIHES Xenopus maternal EST project: interim analysis of the first 13,879 ESTs from unfertilized eggs
Gene 267 (1), 71-87 (2001)
21211403

Contact: Perry J. Blackshear
Office of Clinical Research and Laboratory of Signal Transduction
National Institute of Environmental Health Sciences
A2-05 NIEHS, 101 Alexander Drive, Research Triangle Park, NC 27709,
USA
Tel: 919 541-4899
Fax: 919 541-4571

Email: black009@niehs.nih.gov
Clone is available through Research Genetics, Inc., 2130 Memorial Parkway, Huntsville, AL 35901
phone 800-533-4363 ext.cdna, fax 256-536-9016 att:cdna, email cdna@resgen.com
DNA Sequencing and analyses performed by National Institutes of Health Intramural Sequencing Center (NISC).

PCR Primers
FORWARD: TGTTAAACGACGGCCAGT
BACKWARD: CAGGAACAGCTATGACC
Plate: 0042 row: C column: 09
Seq primer: T7 primer

FEATURES
source
Location/Qualifiers

1..457
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="PBX0042C09"
/clone_lib="Blackshear/Soares normalized Xenopus egg library"
/sex="female"
/tissue_type="unfertilized egg"
/cell_type="unfertilized egg"
/dev_stage="unfertilized egg"
/lab_host="DH10B"
/note="Vector: pT7T3-Pac. Site 1: EcoRI, Site 2: NotI; PolyA-selected mRNA was prepared from unfertilized Xenopus laevis eggs. The library was constructed in the vector pT7T3-Pac as described in Bonaldo, M.F., Lennon, G. and Soares, M.B. 'Normalization and subcloning: two approaches to facilitate gene discovery', Genome Research 6:791-806, 1996. The first strand synthesis used a NotI-dT18 primer: double stranded cDNAs were ligated to EcoRI adapters, digested with NotI, and directionally cloned into the NotI and EcoRI-digested pT7T3-Pac vector. The library contained approximately 7.2 x 10⁵ recombinants, with average insert sizes of 1-1.5 kb."

BASE COUNT 129 a 79 c 175 t
ORIGIN

Query Match 1.8%; Score 172; DB 9; Length 457;
Best Local Similarity 100.0%; Pred. No. 8.9e-69;
Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9290 ACACACTGGTGAATAATGACATTTGCTGATTTTACTGACTGGTCTGTGGGC 9349
|||||
DB 140 ACACACTGGTGAATAATGACATTTGCTGATTTTACTGACTGGTCTGTGGGC 199
|||||
QY 9350 CATCATGCAAGACCATTTAGTGTGCGCAATGTTTCTACTTATTTATGCTGAC 9409
|||||
DB 200 CATCATGCAAGACCATTTAGTGTGCGCAATGTTTCTACTTATTTATGCTGAC 259
|||||
QY 9410 TTGTGAAATGAAATATCTACAAGATAGGCAAGAGTGTCTCTACT 9461
|||||
DB 260 TTGTGAAATGAAATATCTACAAGATAGGCAAGAGTGTCTCTACT 311
|||||

RESULT 9
LOCUS BU069784 669 bp mRNA linear EST 11-DEC-2001
DEFINITION BU069784 NIBB Mochli normalized Xenopus tailbud library Xenopus

ACCESSION laevis cDNA clone XL054124 5', mRNA sequence.
VERSION BU069784
KEYWORDS GI:11798144
SOURCE EST

ORGANISM African clawed frog.
Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae; Xenopodinae; Xenopus.
1 (bases 1 to 669)
Kitayama, A., Terasaka, C., Mochli, M., Ueno, N., Shin-I, T. and Kohara, Y.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Expressed genes in X. laevis embryo
Unpublished (2001)
Contact: Tadao Shin-I
Center For Genetic Resource Information
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp.

FEATURES
source
Location/Qualifiers

1..669
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="XL054124"
/clone_lib="NIBB Mochli normalized Xenopus tailbud library"
/tissue_type="whole embryo"
/dev_stage="stage 25"
BASE COUNT 253 a 122 c 160 g 134 t
ORIGIN

Query Match 0.5%; Score 49; DB 10; Length 669;
Best Local Similarity 100.0%; Pred. No. 2.9e-11;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8171 CGACACAGCGACGACGATACGATGATGTTGTTCCAAAGATTATC 8219
|||||
DB 346 CGACACAGCGACGACGATACGATGATGTTGTTCCAAAGATTATC 394
|||||

RESULT 10
LOCUS AL644957 521 bp mRNA linear EST 07-NOV-2001
DEFINITION AL644957 XGC-egg Silurana tropicalis cDNA clone L1E11f03 5', mRNA sequence.
ACCESSION AL644957
VERSION AL644957.1 GI:16797082
KEYWORDS EST.
SOURCE western clawed frog.
ORGANISM Silurana tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae; Xenopodinae; Silurana.
1 (bases 1 to 521)
Huckle, E., Taylor, R., Ashurst, J.L., Zorn, A.M. and Rogers, J.
Sanger Xenopus tropicalis EST project 2001 (10_2001)
Unpublished (2001)
Contact: Huckle E
Sanger Centre
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS SEQUENCE ID: L1E11f03.sp6
Sequencing primer: Sp6
This sequence is from a Xenopus Gene Collection (XGC) library constructed by Aaron M. Zorn.

FEATURES
source
Location/Qualifiers

1..521
/organism="Silurana tropicalis"
/db_xref="taxon:8364"
/clone="L1E11f03"

```

/clone.lib="XGC-egg"
/dev.stage="egg"
/lab.host="Escherichia coli XL1-blue"
/notes="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
was oligo dt primed from 5ug of poly A+ RNA from egg.
EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end"
BASE COUNT      169 a      97 c      116 g      139 t
ORIGIN

Query Match      0.5%; Score 45; DB 9; Length 521;
Best Local Similarity 100.0%; Pred. No. 2.2e-09;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1538 GAATATGATGACTCTGTTGTACAGAGTTTTCGATTGATGAC 1582
|||||
Db 108 GAATATGATGACTCTGTTGTACAGAGTTTTCGATTGATGAC 152

RESULT 11
AL659719      670 bp      mRNA      linear      EST 13-DEC-2001
LOCUS      AL659719 XGC-neurula Silurana tropicalis cDNA clone TNeu047p06 5',
DEFINITION      mRNA sequence.
ACCESSION      AL659719.1 GI:17673545
VERSION      EST.
KEYWORDS      western clawed frog.
SOURCE      Silurana tropicalis
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Silurana.
1 (bases 1 to 670)
Huckle,E., Taylor,R., Ashurst,J.L., Zorn,A.M. and Rogers,J.
Sanger Xenopus tropicalis EST project 2001 (10_2001)
Unpublished (2001)
Contact: Huckle E
Sanger Centre
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: TNeu047p06.sp6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
Location/Qualifiers
1..670
/organism="Silurana tropicalis"
/db_xref="taxon:8364"
/clone="TNeu047p06"
/clone.lib="XGC-neurula"
/dev.stage="neurula"
/notes="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
was oligo dt primed from 5ug of poly A+ RNA from neurula.
EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end."
BASE COUNT      218 a      158 c      135 g      159 t
ORIGIN

Query Match      0.4%; Score 40; DB 9; Length 670;
Best Local Similarity 100.0%; Pred. No. 4.6e-07;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8328 CACATTATAGAGAATATCATCTTTATCAGGACTAT 8367
|||||
Db 86 CACATTATAGAGAATATCATCTTTATCAGGACTAT 125

RESULT 12
AL632671      648 bp      mRNA      linear      EST 02-NOV-2001
LOCUS      AL632671 XGC-gastrula Silurana tropicalis cDNA clone Tgas020f09 5',
DEFINITION      mRNA sequence.
ACCESSION      AL632671.1 GI:16602154
VERSION      EST.
KEYWORDS      western clawed frog.
SOURCE      Silurana tropicalis
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Silurana.
1 (bases 1 to 648)
Huckle,E., Taylor,R., Ashurst,J.L., Zorn,A.M. and Rogers,J.
Sanger Xenopus tropicalis EST project 2001 (10_2001)
Unpublished (2001)
Contact: Huckle E
Sanger Centre
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: Tgas020f09.sp6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
Location/Qualifiers
1..648
/organism="Silurana tropicalis"
/db_xref="taxon:8364"
/clone="Tgas020f09"
/clone.lib="XGC-gastrula"
/dev.stage="gastrula (stages 10.5-13 mixed)"
/notes="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
was oligo dt primed from 5ug of poly A+ RNA from stages
10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated
into pCS107 with EcoRI at the 5' end and NotI at the 3'
end."
BASE COUNT      252 a      116 c      143 g      134 t      3 others
ORIGIN

Query Match      0.4%; Score 37; DB 9; Length 648;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4426 GTTCAGTGTTCACAGAGCTGTAGTGCATAGAG 4462
|||||
Db 595 GTTCAGTGTTCACAGAGCTGTAGTGCATAGAG 631

RESULT 13
AL638330      681 bp      mRNA      linear      EST 12-DEC-2001
LOCUS      AL638330 XGC-neurula Silurana tropicalis cDNA clone TNeu018h15 5',
DEFINITION      mRNA sequence.
ACCESSION      AL638330.1 GI:16790309
VERSION      EST.
KEYWORDS      western clawed frog.
SOURCE      Silurana tropicalis
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Silurana.
1 (bases 1 to 681)
Huckle,E., Taylor,R., Ashurst,J.L., Zorn,A.M. and Rogers,J.
Sanger Xenopus tropicalis EST project 2001 (10_2001)
Unpublished (2001)
Contact: Huckle E
Sanger Centre
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: TNeu018h15.sp6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
Location/Qualifiers
1..681
/organism="Silurana tropicalis"
/db_xref="taxon:8364"
/clone="TNeu018h15"
/clone.lib="XGC-neurula"
/dev.stage="neurula"
/notes="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
was oligo dt primed from 5ug of poly A+ RNA from neurula.
EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end."
BASE COUNT      218 a      158 c      135 g      159 t
ORIGIN

```

```

FEATURES
  source
    Location/Qualifiers
      1..661
        /organism="Silurana tropicalis"
        /db_xref="taxon:8364"
        /clone="TNeu018h15"
        /clone_id="XGC-neurula"
        /dev_stage="neurula"
        /lab_host="Escherichia coli DH10B"
        /note="Vector: pCS107; Site_1: EcorI; Site_2: NotI; cDNA
        was oligo dt primed from 5ug of poly A+ RNA from neurula.
        EcorI-NotI cut cDNA was then ligated into pCS107 with
        EcorI at the 5' end and NotI at the 3' end."
        EcorI at the 5' end and NotI at the 3' end.
        129 c 158 g 169 t 1 others

BASE COUNT
  224 a 129 c 158 g 169 t 1 others
ORIGIN
  Query Match 0.4%; Score 36; DB 9; Length 661;
  Best Local Similarity 100.0%; Pred. No. 3.4e-05;
  Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 284 GATCGTGTATTATTCACGATCAGACAGTCAA 319
  |||
Db 190 GATCGTGTATTATTCACGATCAGACAGTCAA 225

RESULT 14
  BJO74200 612 bp mRNA linear EST 11-DEC-2001
  LOCUS BJO74200 NIBB Mochli normalized Xenopus tailbud library Xenopus
  DEFINITION laevis cDNA clone XLO92h02 5', mRNA sequence.
  ACCESSION BJO74200
  VERSION BJO74200.1 GI:17504389
  KEYWORDS EST.
  SOURCE African clawed frog.
  ORGANISM Xenopus laevis
  Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
  Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
  Xenopodinae; Xenopus.
  REFERENCE 1 (bases 1 to 612)
  AUTHORS Kitayama,A., Terasaka,C., Mochli,M., Ueno,N., Shin-I,T. and Kohara
  Y.
  TITLE Expressed genes in X. laevis embryo
  JOURNAL Unpublished (2001)
  COMMENT Contact: Tadasu Shin-I
  National Institute of Genetics
  1111 Yata, Mishima, Shizuoka 411-8540, Japan
  Tel: 81-559-81-6856
  Fax: 81-559-81-6855
  Email: tshin@genes.nig.ac.jp.
  Location/Qualifiers
    1..612
      /organism="Xenopus laevis"
      /db_xref="taxon:8355"
      /clone="XLO92h02"
      /clone_id="NIBB Mochli normalized Xenopus tailbud
      library"
      /tissue_type="whole embryo"
      /dev_stage="stage 25"

BASE COUNT
  191 a 144 c 133 g 144 t
ORIGIN
  Query Match 0.3%; Score 33; DB 10; Length 612;
  Best Local Similarity 100.0%; Pred. No. 0.00088;
  Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8330 CATATAGAGAGAAATATCATCATTTATTCACGG 8362
  |||
Db 104 CATATAGAGAGAAATATCATCATTTATTCACGG 136

RESULT 15
  BJO86865/c 533 bp mRNA linear EST 12-DEC-2001
  LOCUS BJO86865

```

```

DEFINITION BJO86865 NIBB Mochli normalized Xenopus tailbud library Xenopus
  laevis cDNA clone XLO54124 3', mRNA sequence.
ACCESSION BJO86865
  BJO86865.1 GI:17583439
VERSION BJO86865.1
  EST.
KEYWORDS EST.
SOURCE African clawed frog.
  ORGANISM Xenopus laevis
  Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
  Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
  Xenopodinae; Xenopus.
  REFERENCE 1 (bases 1 to 533)
  AUTHORS Kitayama,A., Terasaka,C., Mochli,M., Ueno,N., Shin-I,T. and Kohara
  Y.
  TITLE Expressed genes in X. laevis embryo
  JOURNAL Unpublished (2001)
  COMMENT Contact: Tadasu Shin-I
  National Institute of Genetics
  1111 Yata, Mishima, Shizuoka 411-8540, Japan
  Tel: 81-559-81-6856
  Fax: 81-559-81-6855
  Email: tshin@genes.nig.ac.jp.
  Location/Qualifiers
    1..533
      /organism="Xenopus laevis"
      /db_xref="taxon:8355"
      /clone="XLO54124"
      /clone_id="NIBB Mochli normalized Xenopus tailbud
      library"
      /tissue_type="whole embryo"
      /dev_stage="stage 25"

BASE COUNT
  185 a 97 c 85 g 154 t 12 others
ORIGIN
  Query Match 0.3%; Score 28; DB 10; Length 533;
  Best Local Similarity 100.0%; Pred. No. 0.2;
  Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9187 GTCACGCTCTATTATTTATGTTATGTC 9214
  |||
Db 414 GTCACGCTCTATTATTTATGTTATGTC 387

Search completed: October 7, 2002, 13:02:49
Job time : 6584 secs

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 7, 2002, 09:33:18 ; Search time 888 Seconds
(without alignments)
18611.507 Million cell updates/sec

Title: US-09-150-867-2

Perfect score: 9626
Sequence: 1 ggaatccgagatcgatag.....ttaaaaaaaaaacgcaatc 9626

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 1736436 seqs, 858457221 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

N_Geneseq_032802:*

- 1: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	9626	100.0	9626	20	AAK26819	Nucleotide sequenc
2	23	0.2	7684	20	AAK12953	Enterococcus faeca
3	22	0.2	15101	20	ABL08140	Drosophila melanog
4	22	0.2	22191	20	AAK20600	Polynucleotide seq
5	21	0.2	21	20	AAK26825	PCR primer for cen
6	21	0.2	693	23	ABL03005	Drosophila melanog
7	21	0.2	2553	23	ABL26462	Drosophila melanog
8	21	0.2	2751	23	ABL03004	Drosophila melanog
9	21	0.2	3031	23	AAK73906	DNA encoding novel

10	21	0.2	3054	23	AAK74043	DNA encoding novel
C 11	21	0.2	4117	23	ABL27156	Drosophila melanog
C 12	21	0.2	4699	23	ABL12872	Drosophila melanog
C 13	21	0.2	7545	23	AAK73907	DNA encoding novel
14	21	0.2	7971	21	AAK74473	Human immune/haema
15	21	0.2	8617	22	AAK52052	Human polynucleoti
16	21	0.2	9018	24	AAK61167	Human gene regulat
17	21	0.2	10564	16	AAK03854	Human fibroinogen g
18	21	0.2	17141	24	AAK26486	Bovine gamma globl
19	21	0.2	18997	24	ABL32571	Human immune syste
20	20	0.2	754	20	AAK16086	Human gene express
21	20	0.2	951	21	AAK50594	Arabidopsis thalia
22	20	0.2	1143	21	AAK50597	Arabidopsis thalia
23	20	0.2	1198	20	AAK89287	Human regulatory p
24	20	0.2	1202	20	AAK52867	Human prostate tum
25	20	0.2	1205	21	AAK32767	Arabidopsis thalia
26	20	0.2	1215	21	AAK98069	Human colon cancer
27	20	0.2	1440	23	AAK87068	DNA encoding novel
28	20	0.2	2097	14	AAK52316	Coriolus hirsutus
C 29	20	0.2	2336	22	AAK29791	S cerevisiae apopt
30	20	0.2	2391	23	ABL08025	Drosophila melanog
31	20	0.2	2580	22	AAK19182	Human secreted pro
32	20	0.2	3403	23	ABL26975	Drosophila melanog
33	20	0.2	3640	24	AAK99915	Polynucleotide enc
34	20	0.2	4061	22	AAK58478	Human polynucleoti
35	20	0.2	4150	22	AAK51922	Human polynucleoti
C 36	20	0.2	4176	22	AAK52905	Human polynucleoti
C 37	20	0.2	4176	22	AAK52906	Human polynucleoti
C 38	20	0.2	4176	22	AAK52906	Human polynucleoti
39	20	0.2	4231	23	AAK57070	DNA encoding novel
40	20	0.2	4258	22	AAK51921	Human polynucleoti
41	20	0.2	4391	23	ABL08024	Drosophila melanog
42	20	0.2	5875	23	ABL32288	Human immune syste
C 43	20	0.2	5981	24	ABL33762	Human immune syste
44	20	0.2	6012	24	AAK61275	Human gene regulat
45	20	0.2	6222	23	ABL26974	Drosophila melanog

ALIGNMENTS

RESULT 1	AAK26819	
ID	AAK26819	standard; DNA: 9626 BP.
XX	AAK26819:	
AC	XX	
XX	XX	
DT	22-JUN-1999	(first entry)
DE	Nucleotide sequence of centromere-associated protein-E (CENP-E).	
XX	XX	
KW	CENP-E; centromere-associated protein-E; ATPase activity;	
KW	plus end-directed microtubule motor activity; chromosome congression;	
KW	microtubule binding activity; chromosome movement; mitosis;	
KW	cell proliferation; tumor; metastasis; vascular malfunction;	
KW	Inflammatory disease; immune disease; angiogenesis; hypertension;	
KW	stenosis; fungal infection; selective herbicide; fungicide;	
KW	Insecticide; plant growth regulator; activator; cancer cell marker; ss.	
XX	XX	
OS	Xenopus sp.	
XX	XX	
PN	W09913061-A1.	
PD	18-MAR-1999.	
XX	XX	
PF	10-SEP-1998;	98MO-US19231.
XX	XX	
PR	11-SEP-1997;	97US-0058645.
XX	XX	
PA	(REGC) UNIV CALIFORNIA.	
XX	XX	
PI	Cleveland DW, Goldstein LSB, Sakowicz R, Wood KW;	
XX	XX	

DR WPI: 1999-229233/19.
DR P-PSDB: AA01632.

XX Centromere-associated protein-E and related nucleic acid

PS Claim 13; Page 67-73; 77pp: English.

XX The present sequence encodes CENP-E (centromere-associated protein-E)
CC of Xenopus. The protein has at least one of plus end-directed microtubule
CC motor activity, ATPase (adenosine triphosphatase) activity and
CC microtubule binding activity. CENP-E is the motor that powers chromosome
CC movement toward microtubule plus ends and is essential for congression
CC of chromosomes during mitosis. Modulators of CENP-E can thus control
CC cell proliferation. Agents that modulate CENP-E activity are lead
CC therapeutic, biologic, agricultural and diagnostic agents, e.g. for treatment
CC of unwanted cell proliferation (typical of many examples are tumors and
CC metastases; vascular malfunction; inflammatory and immune diseases;
CC angiogenesis; hypertension; restenosis; and fungal infections), also as
CC plant protection agents (selective herbicides, fungicides and
CC insecticides) and plant growth regulators or activators for improving
CC yields. CENP-E is also a diagnostic marker for dividing cells, including
CC cancer cells.

XX Sequence 9626 BP: 3540 A: 1692 C: 2104 G: 2290 T: 0 other:

Query Match 100.0%; Score 9626; DB 20; Length 9626;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 9626; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATTCGGAGTGGATAGGCTACTCGGCGAGGAAATTCAACTGCTTATAGAGAAA 60
DB 1 GAATTCGGAGTGGATAGGCTACTCGGCGAGGAAATTCAACTGCTTATAGAGAAA 60
QY 61 CTGGAACGCCGCCAAAGGAGCTAAAGTGACAGACAGAGGAGCGGTGTCCGATCCGAT 120
DB 61 CTGGAACGCCGCCAAAGGAGCTAAAGTGACAGACAGAGGAGCGGTGTCCGATCCGAT 120
QY 121 TTCTCCACTATAGTGCCTCAAAATGTCGCGAGAGATGCATTAAAGTGTGTGAGGGT 180
DB 121 TTCTCCACTATAGTGCCTCAAAATGTCGCGAGAGATGCATTAAAGTGTGTGAGGGT 180
QY 181 TCGGCGCTTATACAGAGAACAGGGGATCAAGCCATCTGCATGGAAAGGCTGGAAA 240
DB 181 TCGGCGCTTATACAGAGAACAGGGGATCAAGCCATCTGCATGGAAAGGCTGGAAA 240
QY 241 CAACACCATTTCCCAAGTGTGACGACAAAGTCTTCAATTCGATCGTATTTAATTC 300
DB 241 CAACACCATTTCCCAAGTGTGACGACAAAGTCTTCAATTCGATCGTATTTAATTC 300
QY 301 TCACGATCAACAAGTCAAAATTTACCAAGAAATAGCAGTACCTATCATACAGCTTT 360
DB 301 TCACGATCAACAAGTCAAAATTTACCAAGAAATAGCAGTACCTATCATACAGCTTT 360
QY 361 GCAGGATTAATGCGACAATATTTGCAATACGAGACAGACATCTTCAGGACAAACGTACAC 420
DB 361 GCAGGATTAATGCGACAATATTTGCAATACGAGACAGACATCTTCAGGACAAACGTACAC 420
QY 421 AATGATGGAAACCAAAATTCATTTGGGCTATATACCCCAAGCCATACAGAAATTTTAA 480
DB 421 AATGATGGAAACCAAAATTCATTTGGGCTATATACCCCAAGCCATACAGAAATTTTAA 480
QY 481 AATTAATTCAGAGATACCGAACAGAGATTTCTTCAAGAGTTTCTTATATGAGAAATTTA 540
DB 481 AATTAATTCAGAGATACCGAACAGAGATTTCTTCAAGAGTTTCTTATATGAGAAATTTA 540
QY 541 CAATAAATCTGGAAGACCTACTGTGTGATGACAGAAAGAAAGCCCTTGGAATTCG 600
DB 541 CAATAAATCTGGAAGACCTACTGTGTGATGACAGAAAGAAAGCCCTTGGAATTCG 600
QY 601 CGAGGATTTAATAGAAACGTGTATGTCGACCTGCTGAAGAACTTGTAAATGTTCC 660
DB 601 CGAGGATTTAATAGAAACGTGTATGTCGACCTGCTGAAGAACTTGTAAATGTTCC 660

QY 661 TGAACATGTAATACAGTGGATCAAAAAGGTTGAAAAACAGACATTATGAGAGACTAA 720
DB 661 TGAACATGTAATACAGTGGATCAAAAAGGTTGAAAAACAGACATTATGAGAGACTAA 720
QY 721 AATGAATGATCATAGTGTGTTCCACATTAATATTTAGATATGTTGAAAGCCGACA 780
DB 721 AATGAATGATCATAGTGTGTTCCACATTAATATTTAGATATGTTGAAAGCCGACA 780
QY 781 CAGAAATGATCCACAAATTCAGAGAACTGTGATGAGAGCTGTCATGTTTCACACTTGA 840
DB 781 CAGAAATGATCCACAAATTCAGAGAACTGTGATGAGAGCTGTCATGTTTCACACTTGA 840
QY 841 TTGTGTAGATCTGTGCGCAGTGAAGAGACAGCCAACTGAGCTGAGGTGTGAGACT 900
DB 841 TTGTGTAGATCTGTGCGCAGTGAAGAGACAGCCAACTGAGCTGAGGTGTGAGACT 900
QY 901 TAAGGAAGCTGCAACATCAACCGCAGCTGTGTTATCTTGGACGAGTTATTAAGAGCT 960
DB 901 TAAGGAAGCTGCAACATCAACCGCAGCTGTGTTATCTTGGACGAGTTATTAAGAGCT 960
QY 961 TAGCAGCGCCAGGCTGTGATTTTAACTACAGAGACAGCAAACTCACAGAAATCT 1020
DB 961 TAGCAGCGCCAGGCTGTGATTTTAACTACAGAGACAGCAAACTCACAGAAATCT 1020
QY 1021 CCAAAATTCATTTGGGAGGAAATGCTTAAACGCTTATATTTGACAAATTTAGCCAGTTTC 1080
DB 1021 CCAAAATTCATTTGGGAGGAAATGCTTAAACGCTTATATTTGACAAATTTAGCCAGTTTC 1080
QY 1081 TTTTGTAGAGCTTAAGTACACTTCAGTTTCCAGTACTGCAAAACATGTGAGAAATAC 1140
DB 1081 TTTTGTAGAGCTTAAGTACACTTCAGTTTCCAGTACTGCAAAACATGTGAGAAATAC 1140
QY 1141 TCCCATGTTAATGAGTCTCGATGATGATGAGCGTTGCTTAAAAAGTACAGAAAGAAAT 1200
DB 1141 TCCCATGTTAATGAGTCTCGATGATGATGAGCGTTGCTTAAAAAGTACAGAAAGAAAT 1200
QY 1201 CTGTGATTTAAAGAAACATTTAGAGATTTAGATCATGCTGTAACAAAGCTCAAGC 1260
DB 1201 CTGTGATTTAAAGAAACATTTAGAGATTTAGATCATGCTGTAACAAAGCTCAAGC 1260
QY 1261 AATGCTTAAAGAGCATATACACAGTTGCTAGCTGAATCAACACATACAGAAAGAG 1320
DB 1261 AATGCTTAAAGAGCATATACACAGTTGCTAGCTGAATCAACACATACAGAAAGAG 1320
QY 1321 AGAATATGAATATGAGCACTTGCACAAATATGTTGCTTCAATCCCAAGATCTCAACA 1380
DB 1321 AGAATATGAATATGAGCACTTGCACAAATATGTTGCTTCAATCCCAAGATCTCAACA 1380
QY 1381 GGACCAAGGGTCAAAAGCAAAAGAGTACGTGGGCGCAGAGAAATCCAAATAG 1440
DB 1381 GGACCAAGGGTCAAAAGCAAAAGAGTACGTGGGCGCAGAGAAATCCAAATAG 1440
QY 1441 TTTACATGCTTCTGCTGTTCTGACTTTGATATGCTTCCAGATTTACCTGGCAATTTAG 1500
DB 1441 TTTACATGCTTCTGCTGTTCTGACTTTGATATGCTTCCAGATTTACCTGGCAATTTAG 1500
QY 1501 CAAGAAGCAAAAGTCTCTGACATGCTTCATTTCCGAATTTGAGAGCTCTGTTGTAC 1560
DB 1501 CAAGAAGCAAAAGTCTCTGACATGCTTCATTTCCGAATTTGAGAGCTCTGTTGTAC 1560
QY 1561 AGAGTTTCTGATTTGATGACAGCCCTCTCCATGATGAGACAGCAATGGAATGATGACAA 1620
DB 1561 AGAGTTTCTGATTTGATGACAGCCCTCTCCATGATGAGACAGCAATGGAATGATGACAA 1620
QY 1621 ATGGAATTTAGCCAGTAAATTAATCAACAGAGAAAGACATCACTTCAATCAATGAT 1680
DB 1621 ATGGAATTTAGCCAGTAAATTAATCAACAGAGAAAGACATCACTTCAATCAATGAT 1680
QY 1681 AGACTTTGACAGATTTCTGACAGTGTAGTTCAATGATTTCTTAAAGAAACACAGCT 1740
DB 1681 AGACTTTGACAGATTTCTGACAGTGTAGTTCAATGATTTCTTAAAGAAACACAGCT 1740
QY 1741 ACAATACCTCCCAAGACCTGTGATATGGCTGATGACAGAAAGCTTTTGTGAAA 1800

Db	1741																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																												</
----	------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

QY	3961	CTTTGACGAGTCAATTTGAAACCAAGATGAATTTAAGGGCTGCCGAGAAAGCTAAAGAGA	4020
Db	3961	CAATTGAGCGTCAATTTGAAACCAAGATGAATTTAAGGGCTGCCGAGAAAGCTTAAGAGA	4020
QY	4021	ACAGAAAGCAACTGTTGATAGCTTTAGACAACAGCTTTTACATTTGCTGTGAGAAATTTTC	4080
Db	4021	ACAGAAAGCAACTGTTGATAGCTTTAGACAACAGCTTTTACATTTGCTGTGAGAAATTTTC	4080
QY	4081	ATCACCAACCATGATGATGCTTACACAGAAAAAGGTGTCAATTTGGGTGAAGTTAATTC	4140
Db	4081	ATCACCAACCATGATGATGCTTACACAGAAAAAGGTGTCAATTTGGGTGAAGTTAATTC	4140
QY	4141	GTTACAAAGTAAATGCTGCGCTGGTGAAGAGATGAGCTTCAACATCTTGTAGAGCATTT	4200
Db	4141	GTTACAAAGTAAATGCTGCGCTGGTGAAGAGATGAGCTTCAACATCTTGTAGAGCATTT	4200
QY	4201	AGTTTCAGAACTGGAGCTACTTCTGCTCATGTAAAACTGTGGAGGGAGAAAAACCTTGA	4260
Db	4201	AGTTTCAGAACTGGAGCTACTTCTGCTCATGTAAAACTGTGGAGGGAGAAAAACCTTGA	4260
QY	4261	AATCACAAAAAACTCAATGGCTTTGAAAAAGAAATTTGGGCAAAATCTGAAACAAACCGA	4320
Db	4261	AATCACAAAAAACTCAATGGCTTTGAAAAAGAAATTTGGGCAAAATCTGAAACAAACCGA	4320
QY	4321	AGTTTGAAGCTCATGTTTGGAGAAATCTTAAAGGAAGACACATTAAGCTCAAAAGAACAC	4380
Db	4321	AGTTTGAAGCTCATGTTTGGAGAAATCTTAAAGGAAGACACATTAAGCTCAAAAGAACAC	4380
QY	4381	AGAGGAATATTTCTAGTAAAGAAAAATCAATTCAGTTTGAAGAGAGTGTTCAGTGGTTTACA	4440
Db	4381	AGAGGAATATTTCTAGTAAAGAAAAATCAATTCAGTTTGAAGAGAGTGTTCAGTGGTTTACA	4440
QY	4441	GAAGCTGGTGTAGATGAGATAGAGTCCCTGAAAGCACAGCTTAAAGGCGACGAAAGAAAGCT	4500
Db	4441	GAAGCTGGTGTAGATGAGATAGAGTCCCTGAAAGCACAGCTTAAAGGCGACGAAAGAAAGCT	4500
QY	4501	GGAAATTAAGGATAGAGATTAATTTTAACTGTTCACAACTGCAAAACCCAAATTTAGTTGA	4560
Db	4501	GGAAATTAAGGATAGAGATTAATTTTAACTGTTCACAACTGCAAAACCCAAATTTAGTTGA	4560
QY	4561	GGGAAAAATTTGAAACCTCCATTTGCAAGCTACACATGAGGAAGACAGATTTGATGGCGCTTC	4620
Db	4561	GGGAAAAATTTGAAACCTCCATTTGCAAGCTACACATGAGGAAGACAGATTTGATGGCGCTTC	4620
QY	4621	TGGAAGAAATGGAGATTAAGGCTTCTGGAGAAAACTTTGAGCGAAATTCAGATTTTACTGGA	4680
Db	4621	TGGAAGAAATGGAGATTAAGGCTTCTGGAGAAAACTTTGAGCGAAATTCAGATTTTACTGGA	4680
QY	4681	AAGATTTGCAAGAAAGAAAGCTGGAACCTGTCTAACCAACTTGAATTCCTTCAGAAAAAGAT	4740
Db	4681	AAGATTTGCAAGAAAGAAAGCTGGAACCTGTCTAACCAACTTGAATTCCTTCAGAAAAAGAT	4740
QY	4741	GGAGACGTCCTGTTCTATTTAAAGAGTACCTGCAACAGAACTAGAAAGCTTGTGAGTGA	4800
Db	4741	GGAGACGTCCTGTTCTATTTAAAGAGTACCTGCAACAGAAAGCTAGAAAGCTTGTGAGTGA	4800
QY	4801	AAACATTAATTTAAAGAAATATTGACACAAACCTTAAGCATGATTCAGATTACTCAAGC	4860
Db	4801	AAACATTAATTTAAAGAAATATTGACACAAACCTTAAGCATGATTCAGATTACTCAAGC	4860
QY	4861	TCAGCTGCAAGAAAAACAGACGACGCTACAGTGGCTTAAGAAATCTTGCAATCGCTGCTTC	4920
Db	4861	TCAGCTGCAAGAAAAACAGACGACGCTACAGTGGCTTAAGAAATCTTGCAATCGCTGCTTC	4920
QY	4921	TGACAAATTTGCTCAATTAAGTCAAGAAAAAGAAACCTGTCAGATTTGTGTGATCCTCTGGA	4980
Db	4921	TGACAAATTTGCTCAATTAAGTCAAGAAAAAGAAACCTGTCAGATTTGTGTGATCCTCTGGA	4980
QY	4981	AGAAAAGATTTGTTATTTAACTGAAAGATTTGCATCAAAAAAACTTAATGAACAGAAAAAATTT	5040
Db	4981	AGAAAAGATTTGTTATTTAACTGAAAGATTTGCATCAAAAAAACTTAATGAACAGAAAAAATTT	5040

QY	5041	ACTACACTGAAAAGCATGAACTTGAACCAAGCTCAGGTTGAGCTAAAGTCCGAGGTGGAA	5100
Db	5041	ACTACACTGAAAAGAAATGAACCTTGAACCAAGCTCAGGTTGAGCTAAAGTCCGAGGTGGAA	5100
QY	5101	TCTGATGAGAGATATGATGATGATGCAAGGCTCAGTTAGTCCCTTGACGATGAGAAACA	5160
Db	5101	TCTGATGAGAGATATGATGATGATGCAAGGCTCAGTTAGTCCCTTGACGATGAGAAACA	5160
QY	5161	TGATACTGACACACACTTCTTGCTTTAAACAGCAGATGCAAGTATGTTACTCAAGAAA	5220
Db	5161	TGATACTGACACACACTTCTTGCTTTAAACAGCAGATGCAAGTATGTTACTCAAGAAA	5220
QY	5221	GAAAGAGCGCAACAACCCAACTGAAACACTTAAACAGCTGAGTGGACCACTTAAAAAGAA	5280
Db	5221	GAAAGAGCTGCACACAAACCCAACTGAAACACTTAAACAGCTGAGTGGACCACTTAAAAAGAA	5280
QY	5281	TATTGAATGGGTTTGAATTTTAAAAATGAAAGCCAGCAAAAGCACTTAAAGGCAATG	5340
Db	5281	TATTGAATGGGTTTGAATTTTAAAAATGAAAGCCAGCAAAAGCACTTAAAGGCAATG	5340
QY	5341	TCTGTAATATGAGATTAAGAACTTGGAGCAAAAGCCACAGACTTCAATGTAGATAGA	5400
Db	5341	TCTGTAATATGAGATTAAGAACTTGGAGCAAAAGCCACAGACTTCAATGTAGATAGA	5400
QY	5401	AGAGCTTATGAAAAGCTTAAAGATTAAGAGTCAAGCCGTGGAAACTTTAAAGAACTGA	5460
Db	5401	AGAGCTTATGAAAAGCTTAAAGATTAAGAGTCAAGCCGTGGAAACTTTAAAGAACTGA	5460
QY	5461	ACAAAAGTATTTATCTGACCAACAGAAATGAAATGTTATGCTGGAAATGAGAGAAAT	5520
Db	5461	ACAAAAGTATTTATCTGACCAACAGAAATGAAATGTTATGCTGGAAATGAGAGAAAT	5520
QY	5521	GAAAAATAGCCAGAGCACTGTAATGGTGGAGAGGACAGCTGCAGAGAGCTTAAGGA	5580
Db	5521	GAAAAATAGCCAGAGCACTGTAATGGTGGAGAGGACAGCTGCAGAGAGCTTAAGGA	5580
QY	5581	GAGGTGGAATGTCCCTTGAACCTCAAGTGTCTTAAGAAAGCTCAAGAAAGCTTGA	5640
Db	5581	GAGGTGGAATGTCCCTTGAACCTCAAGTGTCTTAAGAAAGGCTCAAGAAAGCTTGA	5640
QY	5641	GCAGCAGAAAAGATTAAGTTGAGAACCTGACGCTCCGAGTTTGTGTCCTGACAGAAAAGT	5700
Db	5641	GCAGCAGAAAAGATTAAGTTGAGAACCTGACGCTCCGAGATTTTGTGTCCTGACAGAAAAGT	5700
QY	5701	CTCTCTTTGGAAAATCAAGATGCTTTATATGTTGCACTGTGAAAAGAACTTAAGGA	5760
Db	5701	CTCTCTTTGGAAAATCAAGATGCTTTATATGTTGCACTGTGAAAAGAACTTAAGGA	5760
QY	5761	AAGAGATGACCTGCAACCAAGTCTAAGCAACAGCTGTCAGAAATTTGAACCTTAACT	5820
Db	5761	AAGAGATGACCTGCAACCAAGTCTAAGCAACAGCTGTCAGAAATTTGAACCTTAACT	5820
QY	5821	GTCTTTAAAAAGAAAGAAATTTGCAATTGGAAACAGAGAGAGTCTTAAGGCTGC	5880
Db	5821	GTCTTTAAAAAGAAAGAAATTTGCAATTGGAAACAGAGAGAGTCTTAAGGCTGC	5880
QY	5881	CAGGAAAACAAATAGATATCACAGAGAAATATCAAAATATAGAAAGACAGTTACTTCAACA	5940
Db	5881	CAGGAAAACAAATAGATATCACAGAGAAATATCAAAATATAGAAAGACAGTTACTTCAACA	5940
QY	5941	AGCCACCAATTTAAAGAAACCTTTATGAAAGAGAGAGTCTTAAGGCTTAAGGAGCA	6000
Db	5941	AGCCACCAATTTAAAGAAACCTTTATGAAAGAGAGTCTTAAGGCTTAAGGAGCA	6000
QY	6001	ACTGCTTTGAAACACAGAAACCTTTAGGAAACATTTGAAGACCAAGACTTGGCATTTGG	6060
Db	6001	ACTGCTTTGAAACACAGAAACCTTTAGGAAACATTTGAAGACCAAGACTTGGCATTTGG	6060
QY	6061	TAAAAATGAGACAGAGAGATGAAAGCTGCCATTAAGATATACCTTTACAGAAAAAGAT	6120
Db	6061	TAAAAATGAGACAGAGAGATGAAAGCTGCCATTAAGATATACCTTTTACAGAAAAAGAT	6120
QY	6121	GTCTTCTCTACAGACACAGATCAATGAAGATGTTACTACTCTAAAGAAAGCTGAGGGTGA	6180

Db 6121 GTCTTCTAGAAAGACAGATCAATGAAATGTACTACTCTAAAGAGGTGAGGGTGA 6180
Qy 6181 AAAAGAACCTTCTACCTTCAGAGACCTTCAAGACAGACTCTTCCCAAGATGAGA 6240
Db 6181 AAAAGAACCTTCTACCTTCAGAGACCTTCAAGACAGACTCTTCCCAAGATGAGA 6240
Qy 6241 GCTCAGAGAGCTTTAAAGACCAAGATTTGCAGTTGGAAAGAGCCGAGAAAGAGATAG 6300
Db 6241 GCTCAGAGAGCTTTAAAGACCAAGATTTGCAGTTGGAAAGAGCCGAGAAAGAGATAG 6300
Qy 6301 TGAAGTACTAATGAATTAAGAATCTCAGTGTAGATCTCTCTCTAGAGAGAGAT 6360
Db 6301 TGAAGTACTAATGAATTAAGAATCTCAGTGTAGATCTCTCTCTAGAGAGAGAT 6360
Qy 6361 TCTTCAGATAGTACGATTTTGAATGAAGCTGTAGCGAAAGGAAAACTTCGCCATTC 6420
Db 6361 TCTTCAGATAGTACGATTTTGAATGAAGCTGTAGCGAAAGGAAAACTTCGCCATTC 6420
Qy 6421 GAAGCAGCACTGGTTTCAGAAATGAGAGCTATCAGTGCATTAAGATAGAGACCA 6480
Db 6421 GAAGCAGCACTGGTTTCAGAAATGAGAGCTATCAGTGCATTAAGATAGAGACCA 6480
Qy 6481 TGCATTTGGCGCAATCTTAAGCGGAAAGATGAAAGCTGTAAATTAATAGCAGTCCGC 6540
Db 6481 TGCATTTGGCGCAATCTTAAGCGGAAAGATGAAAGCTGTAAATTAATAGCAGTCCGC 6540
Qy 6541 TGAAGAAATTAAGATCTGACAAAAGATGATGAAATTCAGAGATTCAAAGAACTCTT 6600
Db 6541 TGAAGAAATTAAGATCTGACAAAAGATGATGAAATTCAGAGATTCAAAGAACTCTT 6600
Qy 6601 GCAGAAACACTCTTCCCATTAAGTGAAGTTATGTACATTAATGAAGCTCAACTCAAT 6660
Db 6601 GCAGAAACACTCTTCCCATTAAGTGAAGTTATGTACATTAATGAAGCTCAACTCAAT 6660
Qy 6661 GCTCAAGCAACAGAAAGAAAGACATCAACAACAACTTCAGAGAAATGAAGAAAGTGA 6720
Db 6661 GCTCAAGCAACAGAAAGAAAGACATCAACAACAACTTCAGAGAAATGAAGAAAGTGA 6720
Qy 6721 TGAAGTATTCGAACACTTATCATCTCTTAAGGAACAGCTGGACCAAAATACAGATGAGCT 6780
Db 6721 TGAAGTATTCGAACACTTATCATCTCTTAAGGAACAGCTGGACCAAAATACAGATGAGCT 6780
Qy 6781 AAGGAATGAAGAACTCAGAAACTATGAACTCTCGGAAAGAAATGATTCATGGAAAAAGA 6840
Db 6781 AAGGAATGAAGAACTCAGAAACTATGAACTCTCGGAAAGAAATGATTCATGGAAAAAGA 6840
Qy 6841 AATCTCAGTCTGGTGTATTAATGAGAAAGAGCCTCAGCAGAAAGATGATGTTCGAGA 6900
Db 6841 AATCTCAGTCTGGTGTATTAATGAGAAAGAGCCTCAGCAGAAAGATGATGTTCGAGA 6900
Qy 6901 ACSTATGATATACTTGAAGAGCAAAACCAAGAAATACAGAGCTGATGGAAAAATCTC 6960
Db 6901 ACSTATGATATACTTGAAGAGCAAAACCAAGAAATACAGAGCTGATGGAAAAATCTC 6960
Qy 6961 CGCTGTATTCAGAGCAACACACTTTCAGCAGAGCTCTCTAGTAGGCTTCAAAAGGA 7020
Db 6961 CGCTGTATTCAGAGCAACACACTTTCAGCAGAGCTCTCTAGTAGGCTTCAAAAGGA 7020
Qy 7021 AACTGAAGCAGACAAACATTCATGATTAATTAAGAGATCTCTGATCCAGAGCTTC 7080
Db 7021 AACTGAAGCAGACAAACATTCATGATTAATTAAGAGATCTCTGATCCAGAGCTTC 7080
Qy 7081 CAGATCCTTTGGAGCTTGCAGAACTGAGCATGTTAAGCTAAATCTCACTGAGAGCCT 7140
Db 7081 CAGATCCTTTGGAGCTTGCAGAACTGAGCATGTTAAGCTAAATCTCACTGAGAGCCT 7140
Qy 7141 TCTGAACAATTTAAGTTTATACCGAAGCTGCTGAGTCAAAAGAGATATAGCTTGAT 7200
Db 7141 TCTGAACAATTTAAGTTTATACCGAAGCTGCTGAGTCAAAAGAGATATAGCTTGAT 7200
Qy 7201 CAAGATTTATGAGAGAGACTTGTCTGAGCAAAAGAGCATGATAGCTGAGACTCA 7260
Db 7201 CAAGATTTATGAGAGAGACTTGTCTGAGCAAAAGAGCATGATAGCTGAGACTCA 7260

Db 7201 CAAGATTTATGAGAGAGACTTGTCTGAGCAAAAGAGCATGATAGCTGAGACTCA 7260
Qy 7261 ACTGCAGTGTTTGGAGAGCAGCAGGAGAGAAATGGTCCGATTTCTCATCTAGAGAACTCA 7320
Db 7261 ACTGCAGTGTTTGGAGAGCAGCAGGAGAGAAATGGTCCGATTTCTCATCTAGAGAACTCA 7320
Qy 7321 GTTCTGTGAATTAATTAATCTTGAATGAGTACTTTTAAAAAGCAATATTAATTCAGAG 7380
Db 7321 GTTCTGTGAATTAATTAATCTTGAATGAGTACTTTTAAAAAGCAATATTAATTCAGAG 7380
Qy 7381 TGTCCAGAGTACTTTTCAGAGGTGACAGTATTTCTTAATCACTAGATCAACATGCA 7440
Db 7381 TGTCCAGAGTACTTTTCAGAGGTGACAGTATTTCTTAATCACTAGATCAACATGCA 7440
Qy 7441 AGAAGACCTTGAGACAAAGAAAGGCTTTATGCAATGGTGTGAGGAAATTTGAGATTCGA 7500
Db 7441 AGAAGACCTTGAGACAAAGAAAGGCTTTATGCAATGGTGTGAGGAAATTTGAGATTCGA 7500
Qy 7501 CGTGCATGCTTAAGAAACTCAGTGAAGGCATGCAACAGGAAATAGGCGCATTTGCTCTAC 7560
Db 7501 CGTGCATGCTTAAGAAACTCAGTGAAGGCATGCAACAGGAAATAGGCGCATTTGCTCTAC 7560
Qy 7561 CATACAGCTCTTAACAAAAGGCTTAAGGCAAGTGTTCAGTGAAGAAATACACCTGAGAT 7620
Db 7561 CATACAGCTCTTAACAAAAGGCTTAAGGCAAGTGTTCAGTGAAGAAATACACCTGAGAT 7620
Qy 7621 AACCGTATATCTGACCAAGTTTGAAGGAAATTCAGAGAGAGAAAGAACAAACAAAGA 7680
Db 7621 AACCGTATATCTGACCAAGTTTGAAGGAAATTCAGAGAGAGAAAGAACAAACAAAGA 7680
Qy 7681 ACTATGCGCGAGATGAGACCAACGCGCCCTAGTGTATGTATGAGAGAAAGAAATGTC 7740
Db 7681 ACTATGCGCGAGATGAGACCAACGCGCCCTAGTGTATGTATGAGAGAAAGAAATGTC 7740
Qy 7741 TAGACTTTTAGGCACTAGTGAAGAACTGTTCAGAGATGAATTCAGAAACTCCAACTCAAGAT 7800
Db 7741 TAGACTTTTAGGCACTAGTGAAGAACTGTTCAGAGATGAATTCAGAAACTCCAACTCAAGAT 7800
Qy 7801 CAAAATCTGAGAAATGAACCTGAACTTGTCAAAAGATGATGCCATGCAAAAGCTGAAA 7860
Db 7801 CAAAATCTGAGAAATGAACCTGAACTTGTCAAAAGATGATGCCATGCAAAAGCTGAAA 7860
Qy 7861 AGTTGCAATTTTGCAGACAAACACTTACAGAGAAAGCGGAAAGCTGAGCTAAAGCGCAT 7920
Db 7861 AGTTGCAATTTTGCAGACAAACACTTACAGAGAAAGCGGAAAGCTGAGCTAAAGCGCAT 7920
Qy 7921 GCAAGTGAACCTTAACCTAAAAAGCAAGATTAATCTTCAGGCTGCAATGAAAGAAATAGAAA 7980
Db 7921 GCAAGTGAACCTTAACCTAAAAAGCAAGATTAATCTTCAGGCTGCAATGAAAGAAATAGAAA 7980
Qy 7981 CCTACAGAAATGCTGCCAAAAGGTGCAATTAATCTTCAGGCTGCAATGAAAGAAATAGAAA 8040
Db 7981 CCTACAGAAATGCTGCCAAAAGGTGCAATTAATCTTCAGGCTGCAATGAAAGAAATAGAAA 8040
Qy 8041 AACTAAGTGGTAAAGATTAAGAAATGAAGAAATTAATTAATCTCAAAACAGACAGCAAGA 8100
Db 8041 AACTAAGTGGTAAAGATTAAGAAATGAAGAAATTAATTAATCTCAAAACAGACAGCAAGA 8100
Qy 8101 GATTGCTTAAAGTCTTGTGTTGAAGATTAAGGAAGAAAGGCTTGGCTAGGTTAAAGAA 8160
Db 8101 GATTGCTTAAAGTCTTGTGTTGAAGATTAAGGAAGAAAGGCTTGGCTAGGTTAAAGAA 8160
Qy 8161 GGAACCTTAGGCGAGACAGCAGACAAAGATTAACACAGTTTGTCTTCAAAAGATTAATCA 8220
Db 8161 GGAACCTTAGGCGAGACAGCAGACAAAGATTAACACAGTTTGTCTTCAAAAGATTAATCA 8220
Qy 8221 GAAAGCTTCAACTTTCCCTGTGACTTGTGGTGGTGAAGGAGATTAATGAGAGAGCAGC 8280
Db 8221 GAAAGCTTCAACTTTCCCTGTGACTTGTGGTGGTGAAGGAGATTAATGAGAGAGCAGC 8280
Qy 8281 AATGCTTGTGCTGAGTGGAGAAAGGCGGCTTGGAAAGGAGAGCTGTACATTTATAGAA 8340
Db 8281 AATGCTTGTGCTGAGTGGAGAAAGGCGGCTTGGAAAGGAGAGCTGTACATTTATAGAA 8340

Best Local Similarity 100.0%; Pred. No. 6.3;
Matches 23: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3973 AATGAAACACAGATGAATTAA 3995
|||||
Db 4287 AATGAAACACAGATGAATTAA 4309

RESULT 3

ABL08140
ID ABL08140 standard; cDNA; 15101 BP.

XX
AC ABL08140;

XX
DT 26-MAR-2002 (first entry)

XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 18902.

XX
KW Drosophila: developmental biology; cell signalling; insecticide;

XX
KW pharmaceutical; gene; ss.

XX
OS Drosophila melanogaster.

XX
PN WO200171042-A2.

XX
PD 27-SEP-2001.

XX
PE 23-MAR-2001; 2001WO-US09231.

XX
PR 23-MAR-2000; 2000US-191637P.

XX
PR 11-JUL-2000; 2000US-0614150.

XX
PA (PEKE) PE CORP NY.

XX
PI Venter JC, Adams M, Li PWD, Myers EW;

XX
DR WPI: 2001-6556860/75.

XX
DR P-PSDB; ABB64037.

XX
PS New isolated nucleic acid detection reagent for detecting 1000 or more

XX
PT genes from Drosophila and for elucidating cell signalling and cell-cell

XX
PT interactions -

XX
XX Claim 1; SEQ ID NO 18902; 21pp + Sequence Listing; English.

XX
XX The invention relates to an isolated nucleic acid detection reagent

XX
CC capable of detecting 1000 or more genes from Drosophila. The invention is

XX
CC useful in developmental biology and in elucidating cell signalling and

XX
CC cell-cell interactions in higher eukaryotes for the development of

XX
CC insecticides, therapeutics and pharmaceutical drugs. The invention

XX
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA

XX
CC sequences (AB101840-AB16175) and the encoded proteins

XX
CC (ABB57737-ABB72072).

XX
CC The sequence data for this patent did not form part of the printed

XX
CC specification, but was obtained in electronic format directly from WIGO

XX
CC at ftp.wigo.int/pub/published_pct_sequences.

XX
XX Sequence 15101 BP; 4366 A; 3330 C; 3468 G; 3937 T; 0 other;

XX
XX Query Match 0.2%; Score 22; DB 23; Length 15101;

XX
XX Best Local Similarity 100.0%; Pred. No. 18;

XX
XX Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX
OY 8784 CATGCAACACGCAAAAGTACA 8805
|||||
Db 13673 CATGCAACACGCAAAAGTACA 13694

XX
RESULT 4
AAX20600/c
ID AAX20600 standard; DNA; 22191 BP.

XX
XX
XX AAX20600;

XX
DT 05-MAY-1999 (first entry)

XX
DE Polynucleotide sequence from the genome of Treponema pallidum.

XX
KW Treponema pallidum infection; syphilis; Borrelia infection; animal;

XX
KW enzyme production; ds.

XX
OS Treponema pallidum.

XX
PN WO9859034-A2.

XX
PD 30-DEC-1998.

XX
PF 23-JUN-1998; 98WO-US13041.

XX
PR 24-JUN-1997; 97US-0050667.

XX
PA (HUMA-) HUMAN GENOME SCI INC.

XX
PI Fraser CM;

XX
DR WPI: 1999-081273/07.

XX
PT New isolated Treponema pallidum nucleic acids - used to develop

XX
PT products for the detection, diagnosis, characterisation, prevention

XX
PT and therapy of T. pallidum infections, particularly syphilis

XX
PS Claim 1; Page 673-685; 1150pp; English.

XX
XX AAX20500-21243 represent polynucleotide sequences from the genome of

XX
CC Treponema pallidum. The sequences can be used for detection.

XX
CC diagnosis, characterisation, prevention and therapy for T. pallidum

XX
CC infections, particularly syphilis. They can also be used for detecting

XX
CC diseases related to Borrelia infections in animals, and for the

XX
CC production of biosynthetic products such as enzymes.

XX
XX Sequence 22191 BP; 4744 A; 5303 C; 6215 G; 5916 T; 13 other;

XX
XX Query Match 0.2%; Score 22; DB 20; Length 22191;

XX
XX Best Local Similarity 100.0%; Pred. No. 18;

XX
XX Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX
OY 5419 AAAGATAAAGAGTCAGCGCTG 5440
|||||

XX
Db 12662 AAAGATAAAGAGTCAGCGCTG 12641

XX
RESULT 5
AAX26825/c

XX
ID AAX26825 standard; DNA; 21 BP.

XX
XX AAX26825;

XX
XX 22-JUN-1999 (first entry)

XX
XX PCR primer for centromere-associated protein-E (CENP-E) DNA.

XX
DE CENP-E; centromere-associated protein-E; ATPase activity;

XX
KW plus end-directed microtubule motor activity; chromosome congression;

XX
KW microtubule binding activity; chromosome movement; mitosis; PCR primer;

XX
KW cell proliferation; tumor; metastasis; vascular malfunction;

XX
KW inflammatory disease; immune disease; angiogenesis; hypertension;

XX
KW restenosis; fungal infection; selective herbicide; fungicide;

XX
KW insecticide; plant growth regulator; activator; cancer cell marker; ss.

XX
XX Synthetic.

XX
OS Xenopus sp.

XX
XX WO9913061-A1.

XX
XX 18-MAR-1999.

XX
XX

PF 10-SEP-1998; 98WO-0519231.
 XX
 PR 11-SEP-1997; 97US-0058645.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Cleveland DW, Goldstein LSB, Sakowicz R, Wood KW;
 XX
 DR WPI; 1999-229233/19.
 XX
 PT Centromere-associated protein-E and related nucleic acid
 XX
 PS Disclosure; Page 44; 77pp; English.
 XX
 CC PCR primers AAX26824-25 were used to amplify nucleic acid encoding
 CC CENP-E (centromere-associated protein-E) of Xenopus. The protein has at
 CC least one of plus end-directed microtubule motor activity, ATPase
 CC (adenosine triphosphatase) activity and microtubule binding activity.
 CC CENP-E is the motor that powers chromosome movement toward microtubule
 CC plus ends and is essential for congression of chromosomes during
 CC mitosis. Modulators of CENP-E can thus control cell proliferation.
 CC Agents that modulate CENP-E activity are lead therapeutic.
 CC cell proliferation and diagnostic agents, e.g. for treatment of unwanted
 CC cell proliferation (typical of many examples are tumors and
 CC angiogenesis; vascular malfunction; inflammatory and immune diseases;
 CC plant-protection agents (selective herbicides, fungicides and
 CC insecticides) and plant growth regulators or activators for improving
 CC yields. CENP-E is also a diagnostic marker for dividing cells, including
 CC cancer cells.
 XX
 SQ Sequence 21 BP; 7 A; 6 C; 5 G; 3 T; 0 other;
 XX
 Query Match 0.2%; Score 21; DB 20; Length 21;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 619 CGGTATGTTGCTGACCTGAC 639
 DB 21 CGGTATGTTGCTGACCTGAC 1
 XX
 RESULT 6
 ABL03005
 ID ABL03005 standard; cDNA; 692 BP.
 XX
 AC ABL03005;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 3497.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 XX
 KW pharmaceutical; gene; ss.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 XX
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 XX
 DR P-PSDB; ABB58902.
 XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 XX
 PS Claim 1; SEQ ID NO 3497; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 692 BP; 201 A; 150 C; 205 G; 136 T; 0 other;
 XX
 Query Match 0.2%; Score 21; DB 23; Length 692;
 Best Local Similarity 100.0%; Pred. No. 51;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 1830 AGTCAAGGAGGAAGAAAAA 1850
 DB 601 AGTCAAGGAGGAAGAAAAA 621
 XX
 RESULT 7
 ABL26462/c
 ID ABL26462 standard; DNA; 2553 BP.
 XX
 AC ABL26462;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 30859.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 XX
 KW pharmaceutical; gene; ds.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 XX
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 XX
 PS Claim 1; SEQ ID NO 30859; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC

CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 2553 BP; 636 A; 561 C; 623 G; 733 T; 0 other;

Query Match 0.2%; Score 21; DB 23; Length 2553;

Best Local Similarity 100.0%; Pred. No. 51;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 5483 CAAGAAATGAAATGGTTATG 5503
|||||

Db 511 CAAGAAATGAAATGGTTATG 491

RESULT 8
ABL03004/C
ID ABL03004 standard; cDNA; 2751 BP.

XX
AC ABL03004;

XX
DT 26-MAR-2002 (first entry)

XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 3494.

XX
KM Drosophila: developmental biology; cell signalling; insecticide;

XX
KW pharmaceutical; gene; ss.

XX
OS Drosophila melanogaster.

XX
PN WO200171042-A2.

XX
PD 27-SEP-2001.

XX
PF 23-MAR-2001; 2001WO-US09231.

XX
PR 23-MAR-2000; 2000US-191637P.

XX
PR 11-JUL-2000; 2000US-0614150.

XX
PA (PEKE) PE CORP NY.

XX
PI Venter JC, Adams M, Li PWD, Myers EW;

XX
DR WPI: 2001-656860/75.

XX
DR P-PSDB; ABB58901.

XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -

XX
PS Claim 1; SEQ ID NO 3494; 21pp + Sequence Listing; English.

XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins
CC (AB57737-AB872072).

XX
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 2751 BP; 748 A; 655 C; 570 G; 778 T; 0 other;

Query Match 0.2%; Score 21; DB 23; Length 2751;

Best Local Similarity 100.0%; Pred. No. 51;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1830 AGTCAAGGAGAGAGAAAAA 1850
|||||

Db 1092 AGTCAAGGAGAGAGAAAAA 1072

RESULT 9
AAS73906
ID AAS73906 standard; cDNA; 3031 BP.

XX
AC AAS73906;

XX
DT 13-FEB-2002 (first entry)

XX
DE DNA encoding novel human diagnostic protein #9710.

XX
KM Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX
OS Homo sapiens.

XX
PN WO200175067-A2.

XX
PD 11-OCT-2001.

XX
PF 30-MAR-2001; 2001WO-US08631.

XX
PR 31-MAR-2000; 2000US-0540217.

XX
PR 23-AUG-2000; 2000US-0649167.

XX
PA (HYSE-) HYSEQ INC.

XX
PI Drmanac RT, Liu C, Tang YT;

XX
DR WPI: 2001-639362/73.

XX
DR P-PSDB; ABC09719.

XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

XX
PS Claim 1; SEQ ID NO 9710; 103pp; English.

XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.

XX
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.

XX
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 3031 BP; 890 A; 707 C; 780 G; 654 T; 0 other;

Query Match 0.2%; Score 21; DB 23; Length 3031;

Best Local Similarity 100.0%; Pred. No. 51;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 7478 TTGAGGAATTGGAGATCTG 7498
|||||

Db 992 TTGAGGAATTGGAGATCTG 1012

```

RESULT 10
AAS74043
ID AAS74043 standard; cDNA; 3054 BP.
XX
XX AAS74043:
XX
XX 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #9847.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX P-PSDB; ABG098356.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity
XX
XX Claim 1; SEQ ID NO 9847; 103bp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
XX CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX CC and gene mapping, and in recombinant production of (II). The
XX CC polynucleotides are also used in diagnostics as expressed sequence tags
XX CC for identifying expressed genes. (I) is useful in gene therapy techniques
XX CC to restore normal activity of (II) or to treat disease states involving
XX CC quantitating a polypeptide in tissue, as molecular weight markers and as
XX CC a food supplement. (II) and its binding partners are useful in medical
XX CC imaging of sites expressing (II). (I) and (II) are useful for treating
XX CC disorders involving aberrant protein expression or biological activity.
XX CC The polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. AAS64197-AAS94564 represent novel human
XX CC diagnostic coding sequences of the invention.
XX CC Note: The sequence data for this patent did not appear in the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 3054 BP; 967 A; 705 C; 712 G; 670 T; 0 other;
XX
Query Match 0.2%; Score 21; DB 23; Length 3054;
Best Local Similarity 100.0%; Pred No. 51;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 7478 TTGGAGGAATTGGAGATCTG 7498
Db 992 TTGAGGGAATTGGAGATCTG 1012

```

```

RESULT 11
ABL27156/c
ABL27156/c

```

```

ID ABL27156 standard; DNA; 4117 BP.
XX
XX ABL27156:
XX
XX 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 33941.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical; gene; ds.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX PT interactions
XX
XX Claim 1; SEQ ID NO 32941; 21bp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX CC useful in developmental biology and in elucidating cell signalling and
XX CC cell-cell interactions in higher eukaryotes for the development of
XX CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX CC discloses genomic DNA sequences (ABL1840-ABL16175) and the encoded proteins
XX CC (ABBS7737-ABB72072).
XX CC The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 4117 BP; 1190 A; 945 C; 975 G; 1007 T; 0 other;
XX
Query Match 0.2%; Score 21; DB 23; Length 4117;
Best Local Similarity 100.0%; Pred No. 52;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 9598 GATTTTAATTAATAAAAAA 9618
Db 704 GATTTTAATTAATAAAAAA 684

```

```

RESULT 12
ABL12872/c
ABL12872/c

```

```

ID ABL12872 standard; cDNA; 4699 BP.
XX
XX ABL12872:
XX
XX 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 33098.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX KW pharmaceutical; gene; ss.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX

```



```
XX 27-SEP-2001.
PD
XX
XX 23-MAR-2001; 2001WO-US09231.
PF
XX
XX 23-MAR-2000; 2000US-191637P.
PR
XX 11-JUL-2000; 2000US-0614150.
PR
XX (PEKE ) PE CORP NY.
PA
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX WPI: 2001-656860/75.
DR
XX P-PSDB; ABB68769.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
PT
XX
XX Claim 1: SEQ ID NO 33098; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (AB57737-AB872072).
CC (AB57737-AB872072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 4699 BP; 1342 A; 1013 C; 1072 G; 1272 T; 0 other:
SQ
Query Match 0.2%; Score 21; DB 23; Length 4699;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 9598 GATTTAAATTAATAAAAAA 9618
DB 3862 GATTTAAATTAATAAAAAA 3842
IIIIIIIIIIIIIIIIIIIIII
RESULT 13
AA573907/C
ID AA573907 standard; cDNA; 7545 BP.
XX
XX AA573907;
AC
XX
XX 13-FEB-2002 (first entry)
DT
XX
XX DNA encoding novel human diagnostic protein #9711.
DE
XX
XX Human: chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200175067-A2.
PN
XX
XX 11-OCT-2001.
PD
XX
XX 30-MAR-2001; 2001WO-US08631.
PF
XX
XX 31-MAR-2000; 2000US-0540217.
PR
XX 23-AUG-2000; 2000US-0649167.
PR
XX (HYSE-) HYSEQ INC.
PA
XX
XX Drmanac RT, Liu C, Tang YT;
PI
XX
XX WPI: 2001-639362/73.
DR
```

```
DR P-PSDB; ABC09720.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
PT
XX
XX Claim 1: SEQ ID NO 9711; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 7545 BP; 1600 A; 1993 C; 1792 G; 2160 T; 0 other:
SQ
Query Match 0.2%; Score 21; DB 23; Length 7545;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 7478 TTGGAGGAAATTGGAGATCTG 7498
DB 6339 TTGGAGGAAATTGGAGATCTG 6319
IIIIIIIIIIIIIIIIIIIIII
RESULT 14
AAK74473
ID AAK74473 standard; DNA; 7971 BP.
XX
XX AAK74473;
AC
XX
XX 07-NOV-2001 (first entry)
DT
XX
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:29285.
DE
XX
XX Human: immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ds.
XX
XX Homo sapiens.
OS
XX
XX WO200157182-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 17-JAN-2001; 2001WO-US01354.
PF
XX
XX 31-JAN-2000; 2000US-0179065.
PR
XX 04-FEB-2000; 2000US-0180628.
PR
XX 24-FEB-2000; 2000US-0184664.
PR
XX 02-MAR-2000; 2000US-0186350.
PR
XX 16-MAR-2000; 2000US-0189874.
PR
XX 17-MAR-2000; 2000US-0190076.
PR
XX 18-APR-2000; 2000US-0198123.
PR
XX 19-MAY-2000; 2000US-0205515.
PR
XX 07-JUN-2000; 2000US-0209467.
PR
XX 28-JUN-2000; 2000US-0214886.
DR
```


CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/haematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/haematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
 CC represent sequences used in the exemplification of the present invention.
 XX

Sequence 7971 BP; 2011 A; 1701 C; 1666 G; 2593 T; 0 other;

Query Match 0.2%; Score 21; DB 22; Length 7971;
 Best Local Similarity 100.0%; Pred. No. 52;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 794 ACAATTTCAGAGAACTGTGAT 814
 ||||||||||||||||||
 DB 2703 ACAATTTCAGAGAACTGTGAT 2723

RESULT 15

AAK52052
 ID AAK52052 standard; cDNA; 8617 BP.

XX AAK52052;

DT 06-NOV-2001 (first entry)

DE Human polynucleotide SEQ ID NO 597.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KM tissue growth factor; immunomodulatory; cancer; leukaemia;
 XX nervous system disorder; arthritis; inflammation; ss.

OS Homo sapiens.

XX WO200157190-A2.

XX 09-AUG-2001.

PF 05-FEB-2001; 2001WO-US04098.

XX 03-FEB-2000; 2000US-0496914.

PR 27-APR-2000; 2000US-0560875.

PR 20-JUN-2000; 2000US-0598075.

PR 19-JUL-2000; 2000US-0620325.

PR 01-SEP-2000; 2000US-0634936.

PR 15-SEP-2000; 2000US-0663561.

PR 20-OCT-2000; 2000US-0693325.

PR 30-NOV-2000; 2000US-0728422.

XX (HSE-) HSEQ INC.

XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;

PI Zhao QA, Wang D, Zhang J, Ren F, Chen R, Wang ZW;

PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX WPI: 2001-476283/51.

DR P-PSDB: AAM78919.

DR Nucleic acids encoding polypeptides with cytokine-like activities,
 PT useful in diagnosis and gene therapy -
 XX
 XX
 PS Claim 1; Page 2144-2154; 6221pp; English.
 XX
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or

CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC actvlin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAM80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.
 XX

Sequence 8617 BP; 2468 A; 2025 C; 2201 G; 1923 T; 0 other;

Query Match 0.2%; Score 21; DB 22; Length 8617;
 Best Local Similarity 100.0%; Pred. No. 52;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 7478 TTGAGGAATTTGAGATCTG 7498
 ||||||||||||||||||
 DB 1207 TTGAGGAATTTGAGATCTG 1227

Search completed: October 7, 2002, 11:17:13
 Job time : 1337 secs

THIS PAGE BLANK (USPTO)

Research, 9500 Gilman Dr., La Jolla, CA 92093-0660, USA
Location/Qualifiers
1. .9610
/organism="Xenopus laevis"
/db_xref="taxon:8335"
1. .9610
/gene="XCENP-E"
135. .8999
/gene="XCENP-E"
/note="Kinetochore-associated; plus end-directed microtubule motor"
/codon_start=1
/product="Kinesin-related protein"
/protein_id="AAC60300.1"
/db_xref="GI:2586071"
/translation="MSBGDAVVCYRVPFLQREGDQANLQMKAGNNTLSQVDTGTS
FNEPVPFNSHESSTOYETIAPVPIIRSLAGYNGTIFVGTSSGKTYTMGTPNSLG
IIPQAIQEFKTIQEIPIRREFLARSYMEIYNEVYKILLCDRRKKPLEIREDENRY
YVADTLELVAVPEHVIQWKKEKNRNGETKMDHSRSHTTFRMIVESRDNDY
NSENCDGAVWVSHNLVDLAGSERASGTGAGVRLKEGCNINSLFTLIGQYIKLSG
QAGGFINRDSKLFRIQNSLGNKTYITCTIPVSGDETLSTLQFASRAKHRTNP
HNVYLDEALLKRYKEIIDLKQLEMLSSSTKQAQAKKEHTVLKOLELKE
REDRIWHLTNIVASVSOESODQYKRRRTVAPDKIQNLSHVSDDMLSTRLE
NRSKAKSDMPSPPEIDSVCTFSDPDDLSDMDNGLDAENNLASKYTHREKSL
HOSMIDRQSDSVQFHDSSKENQIYLPKDSGMACRCRSPFKETITSLQOQSKSE
LEKELVOSFELKLALEEQLSYAKMLNEMVNRREHSINAEVQTEDEKVRKEMV
LDGSGYASNSDIQDSVDQKSSSDECEIHKMLEOKVLEEFTEINNKSEND
KQKSEDFMESIQCEIYNAEKANALEALMAGKNDNMLEETIKREIADLKE
KENQETNEFEILEKETQHEHAOLIHGEGKIKYVENAWNOLPEDLTKRLKE
OEIOLAEKLRADNLOKVRNEDLSVSGMSKEKCEITPOLKQSLDAEVTDQKE
CSFLSENLEKKEMDPSNNYKRNOKRAASLEKQLETKSNKMKDADLOKLOSAF
NEIYINGLAKGVPRLDLSVLEKRVSESKOLEALEKNALENEVYKLSLEYFL
PNEVECLKNQISKASEEIMLKQEGEHSASTISKQETIMOQESQILQITLDEVTHTOS
KVQDTEQYLEMKMHDLFEKYIRNKSEADILREKNEKGMESYEVIAQTKHEL
BETTRDEKQULHEKRYFOAMQITPITPLSDSLPKELVSGNSODPIRENDYHNLIA
LATERNNIWCILETERNSLKEOVIDLNTLOSLQASQIESKLOKPRODESEVYLAE
LEMELKGLHLSQISTEKLQLENTLEYTELQTLQOEMKNTTIERNELQJNEDLAE
HDSLKQDLSNEDISTQDELRRAQOELEKQOLDSVFOQLDSDGSSPNHAY
ANOKSEYLGYNSTOSQMLGEBDELOTSCALVSELELRAHVSVGEENLETKRL
NGELKEKILGSESEVLSKMLENKEDNNLKQAEYSKQFSLDEYFGSQOLV
DELEVALQKLAABERLEIKRDYFELVQANTNLVEGKLETPLOADHEDSDIRSE
EMETIKVGEKLENOYLERLOEKLKSLKLEILOKEMETPLADHEDSDIRSE
ENILKENIDITTLKHSQDQAOLOKTOEOLAKNLAIASNDCPIQTOEKESADCVL
PLEEKILLTEELHOKTNEOEKILLHENELEOAVLEKQVEHLMMSIESKLSL
OHEKHDTQEOULLAKQOMOVYQEKKELOQTHHLRAVDHLEKNEELGVLNENQO
KTKKEOCLLNENKELDSOHRDCEIIEBKSLKDESALETLKESQVYINQDME
MVMHEMELKNSQRTVYAKRDQLODLRESVENSITQDLRAQALQOQKXQVQL
TSQISVLQEKISLENOMLYNAVETTESRDMLQOSKQHFSEIETLSLKEEF
ALQOAEKDKADARKTIDILEKISNIEEOQLQATNIKELTYRESLTIQKQOLANT
EHLRETLKSDIALGKMEORDEAPANKVIALTKMSLSEQLINENVTTLKEBGEKET
FYLOPESKOQSSOMELELRESLTKDLOLEAEKETSAPANEKNTTAKISLSEBITL
ONASILNEAVSERENLRHSKQOLVSELQSLTKRDHAFQSKKEEDINNKLAEVK
AEIKILTKMEDEFROKSELSQSHLSBELQVTELOMKQKQOEKEDINNKLAEVK
EYVDELLQSHLSQEOLOQLOMELRNKLEIELCEMDIMEKEISVLRQMLQNEPQEE
DVAENMDILIESNNOEIOELMEISAVYSEBHNTLSLSLEKLEIETEARHOMLNKE
SLSSTLSRSQSLQTEHVKLNTQLOTLKFKVYVYTAJVEDSHSLKANTQVODSEVO
KRHDELRLOQLOLEOHGRKMSDASELEKCELEFEPLNLEFKANITQVODSEVO
VELNVOGSTLOEILEHKGMOMLEERGLQHDARKLSQMOOPNRISTQILLTKR
LNAVOSKIQORETIVYLNOPEAKLOEKQONKELMRMHGHSVMAEENARLQV
LTKVODESKLOSRIKMLENLVYKDMHNGEKVAILQDLKLSRMAEENARLQV
LTKKQNDQAAKEIENLQKVAQVAPKEIDNLTQVVAIEMEKIKSYATQOQI
AUKSCLEDEKGLRLKELRRQAQNDQITVCPDQKASTPVTCGGGAGIYOST
AMVLVQSEKAALERELSHYKRYKHLSTRMSSEDEBKYSRSEDSSGSHSGSP
HKTEYLRHGPTEPERSEMPSLHSGSPKSESTKRVYSKASDISHSMTSGSKGM
KHLTSPSKGLKRLALSPNBSMPQIHVISPQGLKHLNLTSLFLPDLNLSPPCQOK
VOENLNSPKGLFLDYKSKMPCPSQFPDNLKLGPSLNLTAESNDKSOAEWVYEA
KETAPBCKTS"
misc_difference 1929. .1935
/gene="XCENP-E"
/note="nucleotide sequence is absent in another XCENP-E
cDNA"

Query Match 99.8%; Score 9610; DB 5; Length 9610;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 9610; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 9 GAGTCGGATAGGCTTAGTCGGGCGAGGGAATTCAACTGGTTATAGAAACTTGAACC 68
DB 1 GAGTCGGATAGGCTTAGTCGGGCGAGGGAATTCAACTGGTTATAGAAACTTGAACC 60
QY GCCGCCAAAAGGACTAAAGTGACAGAGACAGGAGGGGTGGTCCGATTTCTCCAC 128
DB GCCGCCAAAAGGACTAAAGTGACAGAGACAGGAGGGGTGGTCCGATTTCTCCAC 120
QY 129 TAATCGCTCTCAAAATGTCCGAGGAGATCAGTTAAAGTGTGTGAGGGTTCGCGCCG 188
DB 121 TAATCGCTCTCAAAATGTCCGAGGAGATCAGTTAAAGTGTGTGAGGGTTCGCGCCG 180
QY 189 TTATACGAGAGAACAGGGGATCAACCCAACTGCATGGAATGGAAGCTGGAACACACCA 248
DB 181 TTATACGAGAGAACAGGGGATCAACCCAACTGCATGGAATGGAAGCTGGAACACACCA 240
QY 249 TTTCCAAAGTTGATGGGCAAAAGCTTTCAATTTGCATGCTGATTTAATTCACAGAA 308
DB 241 TTTCCAAAGTTGATGGGCAAAAGCTTTCAATTTGCATGCTGATTTAATTCACAGAA 300
QY 309 CAACAAGTCAATTTACCAAGAAATAGCACTACCTATCATAGATCAGCTTTCAGAGAT 368
DB 301 CAACAAGTCAATTTACCAAGAAATAGCACTACCTATCATAGATCAGCTTTCAGAGAT 360
QY 369 ATATAGGCAATATTTGATGAGACAGACATCTTCAGGCAAGAGCTGCACAAATGATG 428
DB 361 ATATAGGCAATATTTGATGAGACAGACATCTTCAGGCAAGAGCTGCACAAATGATG 420
QY 429 GAACACCAATTTATTGGGATTAATACCCCAAGCCATACAGAAAGTTTAAATTAATTC 488
DB 421 GAACACCAATTTATTGGGATTAATACCCCAAGCCATACAGAAAGTTTAAATTAATTC 480
QY 489 AGGAGATACGAGACAGAGATTTCTCTAAGATTTCTAATATGAGATTTCAATGAA 548
DB 481 AGGAGATACGAGACAGAGATTTCTCTAAGATTTCTAATATGAGATTTCAATGAA 540
QY 549 CTGTGAAAGACCTACTGTGTGATGACAGAAAGAACCCCTTGGAAATTCGCGAGAT 608
DB 541 CTGTGAAAGACCTACTGTGTGATGACAGAAAGAACCCCTTGGAAATTCGCGAGAT 600
QY 609 TTATAGAAACGTATGTGTGACCTGACCTGAGAACTTGTATGGTCTGTGAACTATG 668
DB 601 TTATAGAAACGTATGTGTGACCTGACCTGAGAACTTGTATGGTCTGTGAACTATG 660
QY 669 TAATACAGTGATCAAAAAGGTGAAAAAACAGACATTATGAGAGACTAAATGATG 728
DB 661 TAATACAGTGATCAAAAAGGTGAAAAAACAGACATTATGAGAGACTAAATGATG 720
QY 729 ATCATAGTGTGCTTACATTAATATTTAGATATGATGTTGAAAGCCGACAGACAATG 788
DB 721 ATCATAGTGTGCTTACATTAATATTTAGATATGATGTTGAAAGCCGACAGACAATG 780
QY 789 ATCCCAAAATTCAGAGAACTGTGATGAGAGCTGATCTCATCTTGAATTTGGTAG 848
DB 781 ATCCCAAAATTCAGAGAACTGTGATGAGAGCTGATCTCATCTTGAATTTGGTAG 840
QY 849 ATCTTGCTGGCGATGAAGAGACAGCCAACTGCAAGCTGAAGGTGTGAGACTTAAGAG 908
DB 841 ATCTTGCTGGCGATGAAGAGACAGCCAACTGCAAGCTGAAGGTGTGAGACTTAAGAG 900
QY 909 GCTGCAACATCAACGCGAGCTTGTATCTTGGACAGGTATTAAGAAGGTTGCGCAG 968
DB 901 GCTGCAACATCAACGCGAGCTTGTATCTTGGACAGGTATTAAGAAGGTTGCGCAG 960
QY 969 GCCAGGCTGTGTGATTTATTAACACTACAGACAGACAACTCACCAGAAATTCACAAAT 1028
DB 961 GCCAGGCTGTGTGATTTATTAACACTACAGACAGACAACTCACCAGAAATTCACAAAT 1020

```
QY 1029 CATGGGAGGAAATGCTAAAAAGGTTATATTTGCAATTTACGCCAGTTCTTTTGATG 1088
|||||
Db 1021 CATGGGAGGAAATGCTAAAAAGGTTATATTTGCAATTTACGCCAGTTCTTTTGATG 1080
QY 1089 AGACTTAAGTACACTTTCAGTTTGGCAGTACTGCCAAACATGTGAGAAATATCTCCCATG 1148
|||||
Db 1081 AGACTTAAGTACACTTTCAGTTTGGCAGTACTGCCAAACATGTGAGAAATATCTCCCATG 1140
QY 1149 TTAATGAGGTCTCGATGATGAAAGCGTTGCTTAAAGGTAACAGAAAGAAATCTTGAT 1208
|||||
Db 1141 TTAATGAGGTCTCGATGATGAAAGCGTTGCTTAAAGGTAACAGAAAGAAATCTTGAT 1200
QY 1209 TTAAGAAACAAATTAGAAATTTAGAGTCAATCGTCTGAACAAAGCTCAAGCAATGGCT 1268
|||||
Db 1201 TTAAGAAACAAATTAGAAATTTAGAGTCAATCGTCTGAACAAAGCTCAAGCAATGGCT 1260
QY 1269 AAGAAGAGCATACAGTGTGCTAGCTGAATCAACACATCACAAAGAGAGAGAT 1328
|||||
Db 1261 AAGAAGAGCATACAGTGTGCTAGCTGAATCAACACATCACAAAGAGAGAGAT 1320
QY 1329 GAATATGCACTTGACAAATATTTGTTGCTTCATCCCAAGAAATCTCAACAGACCATA 1388
|||||
Db 1321 GAATATGCACTTGACAAATATTTGTTGCTTCATCCCAAGAAATCTCAACAGACCATA 1380
QY 1389 GGGTCAACGAAACGAAAGATTACGTGGGCCGACAGAAATCCAAATAGTTTACATG 1448
|||||
Db 1381 GGGTCAACGAAACGAAAGATTACGTGGGCCGACAGAAATCCAAATAGTTTACATG 1440
QY 1449 CTTTGGTGTCTTGACTTGTATGCTATCCGATTACCTGGCAATTTAGACAGAG 1508
|||||
Db 1441 CTTTGGTGTCTTGACTTGTATGCTATCCGATTACCTGGCAATTTAGACAGAG 1500
QY 1509 CAAAGTCTCTGACATGCTTCATTTCCAGAAATTTGATGACTCTGTTTGACAGAGTTT 1568
|||||
Db 1501 CAAAGTCTCTGACATGCTTCATTTCCAGAAATTTGATGACTCTGTTTGACAGAGTTT 1560
QY 1569 CTGATTTTATGAGCCCTCTCATGATGACAGCAATGGAATAGATGCAAGATGGAAT 1628
|||||
Db 1561 CTGATTTTATGAGCCCTCTCATGATGACAGCAATGGAATAGATGCAAGATGGAAT 1620
QY 1629 TAGCCAGTAAATTAATCTCAGAGAAAGAAACATCACTTCATCAATATGATGACTT 1688
|||||
Db 1621 TAGCCAGTAAATTAATCTCAGAGAAAGAAACATCACTTCATCAATATGATGACTT 1680
QY 1689 GACAGATTTCTGACAGTGTTCAGTTTCATGATTTCTTAAGGAAACACAGCTACAAT 1748
|||||
Db 1681 GACAGATTTCTGACAGTGTTCAGTTTCATGATTTCTTAAGGAAACACAGCTACAAT 1740
QY 1749 TCCCCAAAGACTCTGGTATATGGCTGAATGCGAAAGAGCTTCTTTGAAAAAGATCA 1808
|||||
Db 1741 TCCCCAAAGACTCTGGTATATGGCTGAATGCGAAAGAGCTTCTTTGAAAAAGATCA 1800
QY 1809 CAAAGCTCCAGAACAACTACAGTCAAGAGAGAGAAAAAGAACTTGTAAGAGCT 1868
|||||
Db 1801 CAAAGCTCCAGAACAACTACAGTCAAGAGAGAGAAAAAGAACTTGTAAGAGCT 1860
QY 1869 TCGAGCTCAAGATAGCAGAACTGGAAGACAGCTTAGTGTCAAAAGCTTAAATCTAG 1928
|||||
Db 1861 TCGAGCTCAAGATAGCAGAACTGGAAGACAGCTTAGTGTCAAAAGCTTAAATCTAG 1920
QY 1929 TGGTTCACAACTCGAGAGAGCATTCATTAATGCTGAAGTCCAAACAGATGTTGAAG 1988
|||||
Db 1921 TGGTTCACAACTCGAGAGAGCATTCATTAATGCTGAAGTCCAAACAGATGTTGAAG 1980
QY 1989 AAGTTGTGAAAAAGAAATGTCAGTCTTTGAGAGACTGTGTTACATGCTCAAAAGCT 2048
|||||
Db 1981 AAGTTGTGAAAAAGAAATGTCAGTCTTTGAGAGACTGTGTTACATGCTCAAAAGCT 2040
QY 2049 ACCTACAGATAGTTCTGTTGATGTAAGAGCTTAAGACAGCTCCATGATGAGTGTAT 2108
|||||
Db 2041 ACCTACAGATAGTTCTGTTGATGTAAGAGCTTAAGACAGCTCCATGATGAGTGTAT 2100
QY 2109 AACACGAAAAAATGCTGGAACAAAAAGATGTTGATTTAGAAAGTTTATTTGAAACCT 2168
|||||
Db 2101 AACACGAAAAAATGCTGGAACAAAAAGATGTTGATTTAGAAAGTTTATTTGAAACCT 2160
QY 2169 ACAAGAAAGTGAAGATGATTAACAAAAATCTTCTGAGCAAGATTTTATGAGAGAT 2228
|||||
Db 2161 ACAAGAAAGTGAAGATGATTAACAAAAATCTTCTGAGCAAGATTTTATGAGAGAT 2220
QY 2229 AGCTATGTGAAGCTATTAATGAGCAAGAAAGGCAAAATGCACTGAGAGAACTGGCA 2288
|||||
Db 2221 AGCTATGTGAAGCTATTAATGAGCAAGAAAGGCAAAATGCACTGAGAGAACTGGCA 2280
QY 2289 GAGATATTTTGAACAAATATTTCTGAGAGATGAAATCTTAAAAAGGAAATTTGAGAT 2348
|||||
Db 2281 GAGATATTTTGAACAAATATTTCTGAGAGATGAAATCTTAAAAAGGAAATTTGAGAT 2340
QY 2349 TGGAGCTTCACATCAAGAGAAATCAAGAAACCAATGAGTTTGAATTTCTGAGAGAA 2408
|||||
Db 2341 TGGAGCTTCACATCAAGAGAAATCAAGAAACCAATGAGTTTGAATTTCTGAGAGAA 2400
QY 2409 CTCAAAAAGAACAGAGGCAACACTAATCCATGAGATGCGAGTTTAAAGAAATTA 2468
|||||
Db 2401 CTCAAAAAGAACAGAGGCAACACTAATCCATGAGATGCGAGTTTAAAGAAATTA 2460
QY 2469 AAAATGAGAGATGTAACAAATCAAAATCTTGAAGAGATCTAAGAACTTAAACAA 2528
|||||
Db 2461 AAAATGAGAGATGTAACAAATCAAAATCTTGAAGAGATCTAAGAACTTAAACAA 2520
QY 2529 TGAAGAGCAGAAATTCACACTTGACAGAAATTAAGAAACGCGAGATTAACCTGCA 2588
|||||
Db 2521 TGAAGAGCAGAAATTCACACTTGACAGAAATTAAGAAACGCGAGATTAACCTGCA 2580
QY 2589 AAGTACGAAATTTTGAATCTCTCGGTTTCCATGGGTATGATGAGAACTTGTAAGAA 2648
|||||
Db 2581 AAGTACGAAATTTTGAATCTCTCGGTTTCCATGGGTATGATGAGAACTTGTAAGAA 2640
QY 2649 TCTTTCACTGAAGCAATCTTCTTGATGCTGAGCTGAGCTGCGAGTGTGACAGAG 2708
|||||
Db 2641 TCTTTCACTGAAGCAATCTTCTTGATGCTGAGCTGAGCTGCGAGTGTGACAGAG 2700
QY 2709 AATGTTCTTTCTCAGAAAGTAAATCTAGAGCTGAAGAGAAATGAGAGACATCA 2768
|||||
Db 2701 AATGTTCTTTCTCAGAAAGTAAATCTAGAGCTGAAGAGAAATGAGAGACATCA 2760
QY 2769 ACTGATTAATCAAAAAAGAGCTGCTTTGTTGAGAAAGCAGCTGGAACCTGAA 2828
|||||
Db 2761 ACTGATTAATCAAAAAAGAGCTGCTTTGTTGAGAAAGCAGCTGGAACCTGAA 2820
QY 2829 AATCAAACTCAAGAAATGGAAGCTGATTTGCAAGAAAGTTGCAAAAGTCTTTTAT 2888
|||||
Db 2821 AATCAAACTCAAGAAATGGAAGCTGATTTGCAAGAAAGTTGCAAAAGTCTTTTAT 2880
QY 2889 AGATTAATCTTAATAGGCTTCTGACAGAAAGTCCACAGAGATTTGCTTCTG 2948
|||||
Db 2881 AGATTAATCTTAATAGGCTTCTGACAGAAAGTCCACAGAGATTTGCTTCTG 2940
QY 2949 TTTGAATTAAGAAAAAGGTTTCTGAGTTCTCAAAAGCAGCTTGAGAAAGCATTTG 3008
|||||
Db 2941 TTTGAATTAAGAAAAAGGTTTCTGAGTTCTCAAAAGCAGCTTGAGAAAGCATTTG 3000
QY 3009 AAAATGCTTGGAGAAATGAAGTGAATGCTTATCAGAAATCAAAATTTTTCGCAAA 3068
|||||
Db 3001 AAAATGCTTGGAGAAATGAAGTGAATGCTTATCAGAAATCAAAATTTTTCGCAAA 3060
QY 3069 TTTGAATGCTTGAAGAAATCAGATCAGCAAGGCTTCTGAAGAGATTAATGTTTAA 3128
|||||
Db 3061 TTTGAATGCTTGAAGAAATCAGATCAGCAAGGCTTCTGAAGAGATTAATGTTTAA 3120
QY 3129 AAGAGAAACATTTGATCTATTAATTAAGCAAAAGAGATTTATCATGACAGAGCAAT 3188
|||||
Db 3121 AAGAGAAACATTTGATCTATTAATTAAGCAAAAGAGATTTATCATGACAGAGCAAT 3180
QY 3189 AGCAGATTTTACAACTGACTGAGAGAGTGAACACACAGTCAAAAGTCCAGACAGCT 3248
|||||
```

Db 3181 AGCAGATTTTACAAGTGAACGAGTGAACACACACAGTCAAAAGTGCAGCAGACTG 3240
QY 3249 AAGACATATCTTGGAGATGAAGAAATGCATGATCTTTTGGAAAGTATATCGAA 3308
Db 3241 AAGGCAATCTTGGAGATGAAGAAATGCATGATCTTTTGGAAAGTATATCGAA 3300
QY 3309 ACAAAGTGAAGCTGAAGACCTTTTAAAGAAAATGAGAACTTTAAAGGCATATGAGT 3368
Db 3301 ACAAAGTGAAGCTGAAGACCTTTTAAAGAAAATGAGAACTTTAAAGGCATATGAGT 3360
QY 3369 CTGTGGAAGTAAAGATTTGTCAGACAAAACATGATCTGGAAGAACTATTAAGGATTAAG 3428
Db 3361 CTGTGGAAGTAAAGATTTGTCAGACAAAACATGATCTGGAAGAACTATTAAGGATTAAG 3420
QY 3429 AGCAACTGCTTCATGAGAAAAAATCTTTTTCAGCAATGACAGCTATATTTCCGATTA 3488
Db 3421 AGCAACTGCTTCATGAGAAAAAATCTTTTTCAGCAATGACAGCTATATTTCCGATTA 3480
QY 3489 CACCTCTTTAGACTGCTTCCCTCCCTCAAAATTAAGTGAAGGAACTCTCAAGACCCGA 3548
Db 3481 CACCTCTTTAGACTGCTTCCCTCCCTCAAAATTAAGTGAAGGAACTCTCAAGACCCGA 3540
QY 3549 TAGAAATCAATGACTACCAATTTTATAGCCCTTGTCTCAGAAAGAAACAACTATTAAG 3608
Db 3541 TAGAAATCAATGACTACCAATTTTATAGCCCTTGTCTCAGAAAGAAACAACTATTAAG 3600
QY 3609 TGTGCTAGAGACTGAAGAAACAGTCTCAGAGCAAGTATTAATTTGAACACTCAAC 3668
Db 3601 TGTGCTAGAGACTGAAGAAACAGTCTCAGAGCAAGTATTAATTTGAACACTCAAC 3660
QY 3669 TTCAAGTCTTCAAGCACAAGATTTGAAGAAAGTCTGATCTCCAGAAAGCCAAAGCAAGCT 3728
Db 3661 TTCAAGTCTTCAAGCACAAGATTTGAAGAAAGTCTGATCTCCAGAAAGCCAAAGCAAGCT 3720
QY 3729 TGGAAAGAGAGAGTTAAACTGCTTTTGAGATGGAAGTACTGGAAGGACACTTAAGT 3788
Db 3721 TGGAAAGAGAGAGTTAAACTGCTTTTGAGATGGAAGTACTGGAAGGACACTTAAGT 3780
QY 3789 ACTCAGCTGTCTATNAGAAATTTGACAGCTGAGAAATTTGGAAGTTCAGAAAAACTTC 3848
Db 3781 ACTCAGCTGTCTATNAGAAATTTGACAGCTGAGAAATTTGGAAGTTCAGAAAAACTTC 3840
QY 3849 AAACAATTCAGAGAGATGAAGAAATATATCTATGAAAGAAAGTACGCTTCAGCAACT 3908
Db 3841 AAACAATTCAGAGAGATGAAGAAATATATCTATGAAAGAAAGTACGCTTCAGCAACT 3900
QY 3909 TCGAAGACTGAAGCAGAGCATGATGATTAACCAAGACCTTAGTGAACCATTTGAGC 3968
Db 3901 TCGAAGACTGAAGCAGAGCATGATGATTAACCAAGACCTTAGTGAACCATTTGAGC 3960
QY 3969 ACTCAATTTGAACCAAGATGAATTAAGGCTGCCAGAGAAAGCTTAAGAGAACGAAAGC 4028
Db 3961 ACTCAATTTGAACCAAGATGAATTAAGGCTGCCAGAGAAAGCTTAAGAGAACGAAAGC 4020
QY 4029 AACTGTTGATAGCTTTAGACAAACAGCTTTTATGATTTGCTGTAGAGATTTCAATACCA 4088
Db 4021 AACTGTTGATAGCTTTTAGACAAACAGCTTTTATGATTTGCTGTAGAGATTTCAATACCA 4080
QY 4089 ACCATGATGCAATTTCTTAACAGGAAAGGTCTATTGGGTGAAGTTAATTCGTTACAA 4148
Db 4081 ACCATGATGCAATTTCTTAACAGGAAAGGTCTATTGGGTGAAGTTAATTCGTTACAA 4140
QY 4149 GTGAAATGCTGCGTGAAGAAAGATGAGCTTCAACACTCTTTGTAAGCATTAAGTTTCAG 4208
Db 4141 GTGAAATGCTGCGTGAAGAAAGATGAGCTTCAACACTCTTTGTAAGCATTAAGTTTCAG 4200
QY 4209 AACTGAGACTTCTGCTCATGTAAATCTGTGAGAGGAGAAACCTTAATTCACAA 4268
Db 4201 AACTGAGACTTCTGCTCATGTAAATCTGTGAGAGGAGAAACCTTAATTCACAA 4260
QY 4269 AAAAATCTAATGCTTGAAGAAAGATATTGGCAATCTGAAGAAAGCAAGTGTGGA 4328
Db 4261 AAAAATCTAATGCTTGAAGAAAGATATTGGCAATCTGAAGAAAGCAAGTGTGGA 4320

QY 4329 AGTCCATGTTGGAGAACTTAAAGGAAGACACAAATTAAGCTCAAGAACAGCAGGAAT 4388
Db 4321 AGTCCATGTTGGAGAACTTAAAGGAAGACACAAATTAAGCTCAAGAACAGCAGGAAT 4380
QY 4389 ATTCTAGTAAAGAAATCAATTCAGTTTGAAGAGAGTGTGCTGCTCAAGAAAGCTGG 4448
Db 4381 ATTCTAGTAAAGAAATCAATTCAGTTTGAAGAGAGTGTGCTGCTCAAGAAAGCTGG 4440
QY 4449 TAGATGATAGAGTCTGAAAGCAGCTTAAAGGACAGAAAGAAAGGCTGGAATTA 4508
Db 4441 TAGATGATAGAGTCTGAAAGCAGCTTAAAGGACAGAAAGAAAGGCTGGAATTA 4500
QY 4509 AGGATGAGATTAATTTGAACCTTTCAAACTGCAAAACCAATTTACTTGAAGGAAAT 4568
Db 4501 AGGATGAGATTAATTTGAACCTTTGCAAACTGCAAAACCAATTTACTTGAAGGAAAT 4560
QY 4569 TGGAAACTCCATTTGCAAGCTGACATGAGAAAGACAGATTTGATCGGCTTCTGAAGAA 4628
Db 4561 TGGAAACTCCATTTGCAAGCTGACATGAGAAAGACAGATTTGATCGGCTTCTGAAGAA 4620
QY 4629 TGGATATAAAGTCTTGGAGAAAAACTTGAAGCAATCAATTTTACTGGAAGATTCG 4688
Db 4621 TGGATATAAAGTCTTGGAGAAAAACTTGAAGCAATCAATTTTACTGGAAGATTCG 4680
QY 4689 AAGAGAAAGCTGCACTGTCTTAACAACTTGAATCTTCAAAAGAGATGAGAGCT 4748
Db 4681 AAGAGAAAGCTGCACTGTCTTAACAACTTGAATCTTCAAAAGAGATGAGAGCT 4740
QY 4749 CCGTTCTATTAAGATGACCTGCAACAGAGCTAGAAAGCTTCTGAGTGAAGAACTTA 4808
Db 4741 CCGTTCTATTAAGATGACCTGCAACAGAGCTAGAAAGCTTCTGAGTGAAGAACTTA 4800
QY 4809 TTCTAAAGAGATTAATTAACCAACCCCTAAGCAATCTGATGATCTCAAGCTAGCTGC 4868
Db 4801 TTCTAAAGAGATTAATTAACCAACCCCTAAGCAATCTGATGATCTCAAGCTAGCTGC 4860
QY 4869 AGAAAAACAGCAGAGCTAGCAGTTGGCTAAGAAATCTTGCATGCTGCTGCAAT 4928
Db 4861 AGAAAAACAGCAGAGCTAGCAGTTGGCTAAGAAATCTTGCATGCTGCTGCAAT 4920
QY 4929 GTCCATTAACCAAGAAAAAGAAACCTCTGACATTTGTGTCATCTCTGGAAGAAAGA 4988
Db 4921 GTCCATTAACCAAGAAAAAGAAACCTCTGACATTTGTGTCATCTCTGGAAGAAAGA 4980
QY 4989 TATTTATTTAATGAGAAATTTGCATCAAAAACTATTAAGAACGGAATTTCTCATG 5048
Db 4981 TATTTATTTAATGAGAAATTTGCATCAAAAACTATTAAGAACGGAATTTCTCATG 5040
QY 5049 AAAAGAAATGAACCTTGAACAGCTCAGTTGAGCTAAAGTGCAGAGTGAACATCTGATGA 5108
Db 5041 AAAAGAAATGAACCTTGAACAGCTCAGTTGAGCTAAAGTGCAGAGTGAACATCTGATGA 5100
QY 5109 AGATATGATGCAATGCAAGTCTCACTTGAATGCTTGCAGCATGAGAAACATGATACG 5168
Db 5101 AGATATGATGCAATGCAAGTCTCACTTGAATGCTTGCAGCATGAGAAACATGATACG 5160
QY 5169 AACAACTCTTCTGCTTAAACAGAGATGCAAGTGAAGTACTCAAGAAAGAAAGAGAG 5228
Db 5161 AACAACTCTTCTGCTTAAACAGAGATGCAAGTGAAGTACTCAAGAAAGAAAGAGAG 5220
QY 5229 TGCACAAACCCATGAACACTTAAACAGCTGAGTGAACCATCTAAGAAAGAAATTAAT 5288
Db 5221 TGCACAAACCCATGAACACTTAAACAGCTGAGTGAACCATCTAAGAAAGAAATTAAT 5280
QY 5289 TGGGTTTGAATTTTAAAGATGAAGCCGAGCAAAAGCCACTAAAGAGCAATGTCTGA 5348
Db 5281 TGGGTTTGAATTTTAAAGATGAAGCCGAGCAAAAGCCACTAAAGAGCAATGTCTGA 5340
QY 5349 ATGAGATTAAGAACTTGAAGCAAGCCAGACAGATCTCAATGATGATGAAGAGCTTA 5408
Db 5341 ATGAGATTAAGAACTTGAAGCAAGCCAGACAGATCTCAATGATGATGAAGAGCTTA 5400

QY	5409	TGMAAAGCTTAAAGGATTAAGAGCTCAGCGCTGGAAACTTTAAAGATGTGAACAAAG	5468
Db	5401	TGMAAAGCTTAAAGGATTAAGAGCTCAGCGCTGGAAACTTTAAAGATGTGAACAAAG	5460
QY	5469	TAAATTAATCTAACCAAGAAATGGTAATGGTTATGCTGGAAATGGAGAAATTGAATAA	5528
Db	5461	TAAATTAATCTAACCAAGAAATGGTAATGGTTATGCTGGAAATGGAGAAATTGAATAA	5520
QY	5539	GCCAGAGAGCTGTAAATTCGTGAGAGGAGACCAGCTGCAGACGACCTTAAGGAGACTGTTG	5588
Db	5521	GCCAGAGAGCTGTAAATTCGTGAGAGGAGACCAGCTGCAGACGACCTTAAGGAGAGTGTG	5580
QY	5589	AAATGTCATTGAAACTCAGATGATCTTAAGAAAGCTCAAGAGCATTCGACAGCAGA	5648
Db	5581	AAATGTCATTGAAACTCAAGATGATCTTAAGAAAGCTCAAGAGCATTCGACAGCAGA	5640
QY	5649	AAGATTAAGGTCAGAGACTCAGCCCGAGATTTCTGCTCGAGGAAAAAGATCTGCTTC	5708
Db	5641	AAGATTAAGGTCAGAGACTCAGCCCGAGATTTCTGCTCGAGGAAAAAGATCTGCTTC	5700
QY	5709	TGGAATAATCAGATGCTTTATTAATGTTGCAACTGTGAAGAAACTCTAAACGAAAGAGATG	5768
Db	5701	TGGAATAATCAGATGCTTTATTAATGTTGCAACTGTGAAGAAACTCTAAACGAAAGAGATG	5760
QY	5789	ACCTGAAACAGCTCTAAGCAACACCTGTTCTCAGAAATTTAAACTCTGTCTTTAA	5828
Db	5781	ACCTGAAACAGCTCTAAGCAACACCTGTTCTCAGAAATTTAAACTCTGTCTTTAA	5820
QY	5829	AAGAAAAAGAAATTTGCATTGTGSAACAAGCAGAGAGAGCAACACTGATGCTGCCAGAGAA	5888
Db	5821	AAGAAAAAGAAATTTGCATTGTGSAACAAGCAGAGAGAGCAACACTGATGCTGCCAGAGAA	5880
QY	5889	CAATAGATATCACAGAGAAAGATATCAAAATATAGAAACAGTTACTTCAACAAGCCACCA	5948
Db	5881	CAATAGATATCACAGAGAAAGATATCAAAATATAGAAACAGTTACTTCAACAAGCCACCA	5940
QY	5949	ATTTAAAGAAACCTTTATATGAAGAGAGAGTCTTCCAGTGTAAAGCAACTGCGTT	6008
Db	5941	ATTTAAAGAAACCTTTATATGAAGAGAGAGTCTTATCCAGTGTAAAGCAACTGCGTT	6000
QY	6009	TGAACACAGAAACACCTTAAGGAAACATGTGAAGAGCAAAACACTTGGCTTAAATAGG	6068
Db	6001	TGAACACAGAAACACCTTAAGGAAACATGTGAAGAGCAAAACACTTGGCTTAAATAGG	6060
QY	6069	AGCAGAGAGAGATGAAGCTGCCAATTAAGTAATAGCTTTACAGAAAAAGATGCTTCTC	6128
Db	6061	AGCAGAGAGAGATGAAGCTGCCAATTAAGTAATAGCTTTACAGAAAAAGATGCTTCTC	6120
QY	6129	TAGAAGACAGATCAATGAAATGTTACTCTTAAGAGAGGTGAGGTTGAAAAAGAGA	6188
Db	6121	TAGAAGACAGATCAATGAAATGTTACTCTTAAGAGAGGTGAGGTTGAAAAAGAGA	6180
QY	6189	CCCTTCACCTTCAGAGACCTTCAACAGCAGAGCTCTTCCAGATGGGAAGGCTCAAG	6248
Db	6181	CCCTTCACCTTCAGAGACCTTCAACAGCAGAGCTCTTCCAGATGGGAAGGCTCAAG	6240
QY	6249	AGCTTTTAAAGACCAAAAGATTTCAGTGTGGAAGAGGCCAGAGAAGAGATTAAGTGAAGCTA	6308
Db	6241	AGCTTTTAAAGACCAAAAGATTTCAGTGTGGAAGAGGCCAGAGAAGAGATTAAGTGAAGCTA	6300
QY	6309	CTAATGAATTAAGAAATCTCACTGCTTAAGATCTCTTCTCTAGAGAGAGATTTCTTCA	6368
Db	6301	CTAATGAATTAAGAAATCTCACTGCTTAAGATCTCTTCTCTAGAGAGAGATTTCTTCA	6360
QY	6369	ATGCTAGCATTTTGAANTGAAGCTGTAAAGGAAAGGAAACCTTGGCCATTTCGAAGCAGC	6428
Db	6361	ATGCTAGCATTTTGAANTGAAGCTGTAAAGGAAAGGAAACCTTGGCCATTTCGAAGCAGC	6420
QY	6429	AACGTGTTTCAGAAATTTGAGACACCTATCACTGACATTTAAAGAGTGTAGACCACTGATTTG	6488
Db	6421	AACGTGTTTCAGAAATTTGAGACACCTATCACTGACATTTAAAGAGTGTAGACCACTGATTTG	6480
QY	6489	CGCAATCTAAGACGAAAAAGAGTGAAGCTGTAAATTAATTAACCACTGCTGCTGAGAA	6548

Db	6481	CGCAATCTTAAACGGCAAAAGGATGAACTGTGTAATTAATTAATACCCAGTCTCGCTGGAAGAA	6540
Qy	6549	TAAAGATCTTGACAAAGAGATGGATGAATTTCAGATATTCAAAGGAATCTTGCAAGAC	6608
Db	6541	TAAAGATCTCTCAACAAAGAGATGATGAAATTCAGAGATTCMAAGCAATCTTGCAAGAC	6600
Qy	6609	AGCTCTCCCATCTAAGTGAAGAGTTATGTCATATTAAGCTGAACTGCAATGCTCAAGC	6668
Db	6601	AGCTCTCCCATCTAAGTGAAGAGTTATGTCATATTAAGCTGAACTGCAATGCTCAAGC	6660
Qy	6669	AACGAAAGAAGACATCAACAAACCTTGCAAGAAATTTAAGGAAGTGGATGAGCTAT	6728
Db	6661	AACGAAAGAAGACATCAACAAACCTTGCAAGAAATTTAAGGAAGTGGATGAGCTAT	6720
Qy	6729	TGCAACACTTATCATCTCTTAAAGAACAGCTGGACCAATTAAGATGAGCTAAGGAATG	6788
Db	6721	TGCAACACTTATCATCTCTTAAAGAACAGCTGGACCAATTAAGATGAGCTAAGGAATG	6780
Qy	6789	AAAAGCTCAGAAACTATGAACCTCTGCGAAAGAATGGATATCATGGAAGAAAGAAATCTCAG	6848
Db	6781	AAAAGCTCAGAAACTATGAACCTCTGCGAAAGAATGGATATCATGGAAGAAAGAAATCTCAG	6840
Qy	6849	TGCTCGCTTTAATGCACAAAGAGCGTCAGAGAAAGATGATGTCAGAAAGCTATAGG	6908
Db	6841	TGCTCGCTTTAATGCACAAAGAGCGTCAGAGAAAGATGATGTCAGAAAGCTATAGG	6900
Qy	6909	ATATACCTTGAGAGAGCAAGAACCAAGAAATACAGAGCTGATGGAAGAAAGAAATCTCCGCTGCT	6968
Db	6901	ATATACCTTGAGAGAGCAAGAACCAAGAAATACAGAGCTGATGGAAGAAAGAAATCTCCGCTGCT	6960
Qy	6969	ATTCAGAGCAACACACTTTCCTCAGCAGTCTCTATGAGAGCTTCAAAAGAAACTGAAG	7028
Db	6961	ATTCAGAGCAACACACTTTCCTCAGCAGTCTCTATGAGAGCTTCAAAAGAAACTGAAG	7020
Qy	7029	CACACAAACATTCGATGTTAAATATAAGGAATCTCTGATCACTCAAGCGTCTCCAGATCCT	7088
Db	7021	CACACAAACATTCGATGTTAAATATAAGGAATCTCTGATCACTCAAGCGTCTCCAGATCCT	7080
Qy	7089	TTGGCAGCTTGCAACACAGAGCATGTTAAGCTAAATACTCACTGACAGACCCCTTGCAACA	7148
Db	7081	TTGGCAGCTTGCAACACAGAGCATGTTAAGCTAAATACTCACTGACAGACCCCTTGCAACA	7140
Qy	7149	AAATTTAAGGTTGTATACCGAATCTGTCAGTCAAGAGATCATATAGCTTATCAAGAAAT	7208
Db	7141	AAATTTAAGGTTGTATACCGAATCTGTCAGTCAAGAGATCATATAGCTTATCAAGAAAT	7200
Qy	7209	ATGAGAAAGGACCTTGCTGCTGAGCAAAAGAGCATATAGCTGGGACTCCAACTGCAGT	7268
Db	7201	ATGAGAAAGGACCTTGCTGCTGAGCAAAAGAGCATATAGCTGGGACTCCAACTGCAGT	7260
Qy	7269	GTTTGAGACACAGCGGCAAGAAATGTCGGATTTCTGCATCTGAGCAACTCAAGTTCTGTG	7328
Db	7261	GTTTGAGACACAGCGGCAAGAAATGTCGGATTTCTGCATCTGAGCAACTCAAGTTCTGTG	7320
Qy	7329	AAATTTGAATTTCTGTAATGAGTTACTTTTAAAAAGCAATATTAATTGAGAGTCTCAGG	7388
Db	7321	AAATTTGAATTTCTGTAATGAGTTACTTTTAAAAAGCAATATTAATTGAGAGTCTCAGG	7380
Qy	7389	ATGACTTTTCAGAGGTCAGAGTATTCCTTAATCACTAGGATCAACACTGCAAGAGAGC	7448
Db	7381	ATGACTTTTCAGAGGTCAGAGTATTCCTTAATCACTAGGATCAACACTGCAAGAGAGC	7440
Qy	7449	TTGAGCAACAAGAAAGGCTTTATGCACTGGTGGAGAAATTTGGAATCTGCAAGCTCAATG	7508
Db	7441	TTGAGCAACAAGAAAGGCTTTATGCACTGGTGGAGAAATTTGGAATCTGCAAGCTCAATG	7500
Qy	7509	CTAAGAAACTCAGTGAAGGCGATCAACACAGAAATATAGCGCATTTGCTTCTACCATACAGC	7568
Db	7501	CTAAGAAACTCAGTGAAGGCGATCAACACAGAAATATAGCGCATTTGCTTCTACCATACAGC	7560
Qy	7569	TCCTTACAAAAAGGCTAAGGCAAGTTGTTCACTCGAAATATACACGTCAGATTAACGTTAT	7628

```
Db 7561 TCTTAACAAAAAGGCTAAAGGAGTGTTCAGTCGAAAAATACACGTGAGATTAACCCGTAT 7620
Qy 7629 ATCTGAACCAAGTTTGAAGGCAATTTGCAAGAGAGAAAGACAAAAACAAGACTTATTC 7688
Db 7621 ATCTGAACCAAGTTTGAAGGCAATTTGCAAGAGAGAAAGACAAAAACAAGACTTATTC 7680
Qy 7689 GCAGATGAGAGCACCAGGCGCTAGTGTAGTAAATGAGAGAAATGCTTAACCTT 7748
Db 7681 GCAGATGAGAGCACCAGGCGCTAGTGTAGTAAATGAGAGAAATGCTTAACCTT 7740
Qy 7749 TAGGCTACTGAAAACCTGTTCAAGATGAATCCAGAAAATCCCAATCAAGATCAAAATGC 7808
Db 7741 TAGGCTACTGAAAACCTGTTCAAGATGAATCCAGAAAATCCCAATCAAGATCAAAATGC 7800
Qy 7809 TAGAAAATCACTGAACCTGTGTCAGAAAGATGATCCATGACACAAAGGTGCAAAAGTTGCA 7868
Db 7801 TAGAAAATCACTGAACCTGTGTCAGAAAGATGATCCATGACACAAAGGTGCAAAAGTTGCA 7860
Qy 7869 TTTTGCAGACAACTACTAAGCAGAAACGCCAGCTGAGCTTAAGCCATGAGAGTGA 7928
Db 7861 TTTTGCAGACAACTACTAAGCAGAAACGCCAGCTGAGCTTAAGCCATGAGAGTGA 7920
Qy 7929 AACTACTAAAAAGCAAAATATCTTCAGCTGCATGAAAAAGAAATAGAAAACCTAGAGA 7988
Db 7921 AACTACTAAAAAGCAAAATATCTTCAGCTGCATGAAAAAGAAATAGAAAACCTAGAGA 7980
Qy 7989 AAATGCTTGCAGAAAGCTCAGTACCATTAAGAAGAAATGACAACTTAAACTAAG 8048
Db 7981 AAATGCTTGCAGAAAGCTCAGTACCATTAAGAAGAAATGACAACTTAAACTAAG 8040
Qy 8049 TGTGTAAGATTGAATGAAAAAATTAAGTACTCAAAAAGCAACAGACCAAGATTCCT 8108
Db 8041 TGTGTAAGATTGAATGAAAAAATTAAGTACTCAAAAAGCAACAGACCAAGATTCCT 8100
Qy 8109 ACTTAAAGTCTTGTGGAAGATGAAGAAAGGCTTGCCTAGTAAAGAGAACCTTA 8168
Db 8101 ACTTAAAGTCTTGTGGAAGATGAAGAAAGGCTTGCCTAGTAAAGAGAACCTTA 8160
Qy 8169 GGGCAGACAGGACAGACAGATCAACAGTTTGTCTCCAAAAGATTAACAGAAAGCTT 8228
Db 8161 GGGCAGACAGGACAGACAGATCAACAGTTTGTCTCCAAAAGATTAACAGAAAGCTT 8220
Qy 8229 CAACCTTCCCTGTGACTTGTGTGTGTGAGGCGGTATAGTGCAGAGACAGCAATGCTTG 8288
Db 8221 CAACCTTCCCTGTGACTTGTGTGTGTGAGGCGGTATAGTGCAGAGACAGCAATGCTTG 8280
Qy 8289 TGTCTGAGTGGAAAAAGCGGCTTGGAAAAAGGAGCTGTCACTTATTAAGAAATAATATC 8348
Db 8281 TGTCTGAGTGGAAAAAGCGGCTTGGAAAAAGGAGCTGTCACTTATTAAGAAATAATATC 8340
Qy 8349 ATCATTTATCAGGACATATGTCAAGTTCTGAAGATGAAAAGAAAGCAAAAGCTTG 8408
Db 8341 ATCATTTATCAGGACATATGTCAAGTTCTGAAGATGAAAAGAAAGCAAAAGCTTG 8400
Qy 8409 ATGCTCATCTCTTCATCTACTGATCATCACACAGAGGCTGCGCTCAAAAACCTGAAACTT 8468
Db 8401 ATGCTCATCTCTTCATCTACTGATCATCACACAGAGGCTGCGCTCAAAAACCTGAAACTT 8460
Qy 8469 ACAGAAATGGCCCTGTTACTCCAGAAAGTCTGAAAATGCCAAAGCTTACACTAGATCTC 8528
Db 8461 ACAGAAATGGCCCTGTTACTCCAGAAAGTCTGAAAATGCCAAAGCTTACACTAGATCTC 8520
Qy 8529 CAAAGAAAGTCGAGATCCACACTAAACGTTGTGTCTCACCACAAAGGTCGAAATCTACA 8588
Db 8521 CAAAGAAAGTCGAGATCCACACTAAACGTTGTGTCTCACCACAAAGGTCGAAATCTACA 8580
Qy 8589 GCCAATTTAGTAATGTCTCCAGGCAAGACCGGATGCATTAACATATACCTTCTCCAGCA 8648
Db 8581 GCCAATTTAGTAATGTCTCCAGGCAAGACCGGATGCATTAACATATACCTTCTCCAGCA 8640
Qy 8649 AGGTGGAATGCAAAAAAGGCTGCCCTGTCTCCAAACAGATGAGATGCCCAACCCAGC 8708
Db 8641 AGGTGGAATGCAAAAAAGGCTGCCCTGTCTCCAAACAGATGAGATGCCCAACCCAGC 8700

Qy 8709 ATGTCATATCCCTGGCAGACCGGACTGCATTAATAATCTAAGTGAAGACGCTTATTCG 8768
Db 8701 ATGTCATATCCCTGGCAGACCGGACTGCATTAATAATCTAAGTGAAGACGCTTATTCG 8760
Qy 8769 ACAATTTGTCTTCCATGCAACAGCAAAAGATACAGAAAAATCTAAATTTCCCTTAAG 8828
Db 8761 ACAATTTGTCTTCCATGCAACAGCAAAAGATACAGAAAAATCTAAATTTCCCTTAAG 8820
Qy 8829 GCAATATTTATGATGTAAGTCAAAATCGATGCCCTTACTGTCATCTCAGTGTGATTA 8888
Db 8821 GCAATATTTATGATGTAAGTCAAAATCGATGCCCTTACTGTCATCTCAGTGTGATTA 8880
Qy 8889 ATTCTAAGCTTGTGATTTTTCAGAGCTCAACAGACAGAGACATGACAAAAGTCAG 8948
Db 8881 ATTCTAAGCTTGTGATTTTTCAGAGCTCAACAGACAGAGACATGACAAAAGTCAG 8940
Qy 8949 CTGAGAACTGTGTATGAGCAAAAAAGAAACGGCACCTGAATGTAAACATCTTGA 9008
Db 8941 CTGAGAACTGTGTATGAGCAAAAAAGAAACGGCACCTGAATGTAAACATCTTGA 9000
Qy 9009 TCCCTGACATCTGACTCTCCTGCTGCAAAAGAGACTGCTACTCTGCTTCTTGAAG 9068
Db 9001 TCCCTGACATCTGACTCTCCTGCTGCAAAAGAGACTGCTACTCTGCTTCTTGAAG 9060
Qy 9069 AAGAAACACTAGAAACTGCCATGTCTCATPAAAGAGAGTCTCACTGGAAGCAAAAGTTGTT 9128
Db 9061 AAGAAACACTAGAAACTGCCATGTCTCATPAAAGAGAGTCTCACTGGAAGCAAAAGTTGTT 9120
Qy 9129 CTTTACTATATCACTGCTTGTGGGAGTGTGTAGCTCTTTAAATTAAGTCAATACGT 9188
Db 9121 CTTTACTATATCACTGCTTGTGGGAGTGTGTAGCTCTTTAAATTAAGTCAATACGT 9180
Qy 9189 CTCAGCTATTTTATATGTAATGTCTGTGTATTTGTTTACACTTTAAGTCCCTGACTT 9248
Db 9181 CTCAGCTATTTTATATGTAATGTCTGTGTATTTGTTTACACTTTTAAAGTCCCTGACTT 9240
Qy 9249 CATATTTGGCTATCTGTATGTTCTTTGTGTGCGCACACACACACTGTGAAAAATGA 9308
Db 9241 CATATTTGGCTATCTGTATGTTCTTTGTGTGCGCACACACACACTGTGAAAAATGA 9300
Qy 9309 CATTTGCAATGATTTTACTGACTGCTGCTGCTGCGGCAATGCAAAACACCAT 9368
Db 9301 CATTTGCAATGATTTTACTGACTGCTGCTGCTGCGGCAATGCAAAACACCAT 9360
Qy 9369 TAGTGTCCAAATGTTTCTACTACTTATTTATATGTCTGACTTTGTGAATAGAAATAC 9428
Db 9361 TAGTGTCCAAATGTTTCTACTACTTATTTATATGTCTGACTTTGTGAATAGAAATAC 9420
Qy 9429 TACAAAAGATGAGGCAAAAGTTGTTTCTCACTAGTAAAGATATGGGTTTAACCTTAAT 9488
Db 9421 TACAAAAGATGAGGCAAAAGTTGTTTCTCACTAGTAAAGATATGGGTTTAACCTTAAT 9480
Qy 9489 TATTCCTCTAATTAATATCTATATATATGATGATAGTATAGTATGATGATG 9548
Db 9481 TATTCCTCTAATTAATATCTATATATATGATGATAGTATAGTATGATGATG 9540
Qy 9549 TGCACCTGTGAAACCTTTTAAATTTCACTGTAATTAAGACGCCGTGATTTAAAT 9608
Db 9541 TGCACCTGTGAAACCTTTTAAATTTCACTGTAATTAAGACGCCGTGATTTAAAT 9600
Qy 9609 AAAAAAAAAA 9618
Db 9601 AAAAAAAAAA 9610

RESULT 2
AB011407 431 bp mRNA linear BCT 27-JAN-1999
LOCUS Buchnera aphidicola mRNA for riboflavin synthase beta chain,
DEFINITION partial cds.
ACCESSION AB011407
VERSION AB011407.1 GI:4190955
```

```

KEYWORDS      riboflavin synthase beta chain.
SOURCE         Buchnera aphidicola (specific_host:Acyrthosiphon pisum) cDNA to
               mRNA.
ORGANISM       Buchnera aphidicola
REFERENCE      Bacteria: Proteobacteria; gamma subdivision; Buchnera.
AUTHORS        1 (bases 1 to 431)
TITLE          Nakabachi, A.
JOURNAL        Direct Submission
               Submitted (21-FEB-1998) Atsushi Nakabachi, University of Tokyo,
               Department of Biological Sciences, Hongo, Bunkyo-ku, Tokyo
               113-0033, Japan (E-mail: ss7722@hongo.ecc.u-tokyo.ac.jp,
               Tel:+81-3-3812-2111, Fax:+81-3-3816-1965)
REFERENCE      2 (sites)
AUTHORS        Nakabachi, A. and Ishikawa, H.
TITLE          Provision of riboflavin to the host aphid, Acyrthosiphon pisum, by
               endosymbiotic bacteria, Buchnera
JOURNAL        J. Insect Physiol. 45, 1-6 (1999)
FEATURES       Location/Qualifiers
               1..431
               /organism="Buchnera aphidicola"
               /specific_host="Acyrthosiphon pisum"
               /db_xref="taxon:9"
               <1..354
               /note="probable"
               /codon_start=1
               /transl_table=11
               /product="riboflavin synthase beta chain"
               /protein_id="BAA74435.1"
               /db_xref="GI:4190956"
               /translation="VHEENILKIVYGFVEIPTIASYIAKSGYDAIIAGITIKGOT
               DHFKYIANDTSSLSRISNQYLPITIGILITKNIKESIGSTKMGKNGSDAALAL
               EMINVMKIKKIVYY"
BASE COUNT    182 a      58 c      59 g      132 t
ORIGIN
Query Match   0.2%; Score 24; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3957 AAAACATGAGCAGTCATGATGAAA 3980
Db 230 AAAACATGAGCAGTCATGATGAAA 253

RESULT 3
LOCUS          HUMTCBXN
DEFINITION     Human T-cell receptor aberrantly rearranged beta-chain V1-D2.1
               gene.
ACCESSION      M13575.1 GI:338940
VERSION        M13575.1
KEYWORDS       D-region; T-cell receptor; T-cell receptor beta-chain; V-region;
               processed gene; pseudogene.
SOURCE         Homo sapiens DNA.
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
               1 (bases 1 to 495)
               Duby, A.D. and Seidman, J.G.
               Abnormal recombination products result from aberrant DNA
               rearrangement of the human T-cell antigen receptor beta-chain gene
               Proc. Natl. Acad. Sci. U.S.A. 83 (13), 4890-4894 (1986)
JOURNAL        86259700
MEDLINE
COMMENT        The coding sequence of D2.1 has been deleted but its 5' flanking
               sequence was retained. This V segment appears to be a pseudogene
               because it has a nucleotide deletion at position 284 resulting in a
               frameshift mutation. Between V1 and D2.1 there are 19 nucleotides
               of unknown origin.
FEATURES       Location/Qualifiers
               1..495
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /cell_line="CEM"

KEYWORDS      riboflavin synthase beta chain.
SOURCE         Buchnera aphidicola (specific_host:Acyrthosiphon pisum) cDNA to
               mRNA.
ORGANISM       Buchnera aphidicola
REFERENCE      Bacteria: Proteobacteria; gamma subdivision; Buchnera.
AUTHORS        1 (bases 1 to 431)
TITLE          Nakabachi, A.
JOURNAL        Direct Submission
               Submitted (21-FEB-1998) Atsushi Nakabachi, University of Tokyo,
               Department of Biological Sciences, Hongo, Bunkyo-ku, Tokyo
               113-0033, Japan (E-mail: ss7722@hongo.ecc.u-tokyo.ac.jp,
               Tel:+81-3-3812-2111, Fax:+81-3-3816-1965)
REFERENCE      2 (sites)
AUTHORS        Nakabachi, A. and Ishikawa, H.
TITLE          Provision of riboflavin to the host aphid, Acyrthosiphon pisum, by
               endosymbiotic bacteria, Buchnera
JOURNAL        J. Insect Physiol. 45, 1-6 (1999)
FEATURES       Location/Qualifiers
               1..431
               /organism="Buchnera aphidicola"
               /specific_host="Acyrthosiphon pisum"
               /db_xref="taxon:9"
               <1..354
               /note="probable"
               /codon_start=1
               /transl_table=11
               /product="riboflavin synthase beta chain"
               /protein_id="BAA74435.1"
               /db_xref="GI:4190956"
               /translation="VHEENILKIVYGFVEIPTIASYIAKSGYDAIIAGITIKGOT
               DHFKYIANDTSSLSRISNQYLPITIGILITKNIKESIGSTKMGKNGSDAALAL
               EMINVMKIKKIVYY"
BASE COUNT    182 a      58 c      59 g      132 t
ORIGIN
Query Match   0.2%; Score 24; DB 1; Length 431;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3957 AAAACATGAGCAGTCATGATGAAA 3980
Db 230 AAAACATGAGCAGTCATGATGAAA 253

RESULT 3
LOCUS          HUMTCBXN
DEFINITION     Human T-cell receptor aberrantly rearranged beta-chain V1-D2.1
               gene.
ACCESSION      M13575.1 GI:338940
VERSION        M13575.1
KEYWORDS       D-region; T-cell receptor; T-cell receptor beta-chain; V-region;
               processed gene; pseudogene.
SOURCE         Homo sapiens DNA.
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
               1 (bases 1 to 495)
               Duby, A.D. and Seidman, J.G.
               Abnormal recombination products result from aberrant DNA
               rearrangement of the human T-cell antigen receptor beta-chain gene
               Proc. Natl. Acad. Sci. U.S.A. 83 (13), 4890-4894 (1986)
JOURNAL        86259700
MEDLINE
COMMENT        The coding sequence of D2.1 has been deleted but its 5' flanking
               sequence was retained. This V segment appears to be a pseudogene
               because it has a nucleotide deletion at position 284 resulting in a
               frameshift mutation. Between V1 and D2.1 there are 19 nucleotides
               of unknown origin.
FEATURES       Location/Qualifiers
               1..495
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /cell_line="CEM"

gene          /cell_type="T cell"
              join(15..63,152..>495)
              /gene="pseudo-TCB"
              join(15..63,152..>495)
              /gene="pseudo-TCB"
              /codon_start=1
              /pseudo
              64..151
              /number=1
              152..>495
              /gene="pseudo-TCB"
              /number=2
              441..442
              /gene="pseudo-TCB"
              /organism="Homo sapiens"
              460..461
              /gene="pseudo-TCB"
              /organism="Homo sapiens"
              103 a      135 c      132 g      124 t      1 others
BASE COUNT
ORIGIN
Query Match   0.2%; Score 24; DB 9; Length 495;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5855 GCAGAGAGAGCAAGCTGATGCT 5878
Db 149 GCAGAGAGAGCAAGCTGATGCT 126

RESULT 4
LOCUS          HUMTCBZE
DEFINITION     Human T-cell receptor unproductively rearranged beta chain gene
               V-region.
ACCESSION      M31347.1 GI:339021
VERSION        M31347.1
KEYWORDS       T-cell receptor; V-region; processed gene; pseudogene.
SOURCE         Human T-cell line HPB-ALL DNA, clone lambda HPB-2.
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
               1 (bases 1 to 500)
               Ikuta, K., Ogura, T., Shimizu, A. and Honjo, T.
               A joining-diversity-joining complex generated by inversion
               mechanism and a variable-diversity complex in the beta-chain gene
               of the human T-cell receptor
               Nucleic Acids Res. 14 (12), 4899-4909 (1986)
JOURNAL        86259066
MEDLINE
COMMENT        Location/Qualifiers
               1..500
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /map="7q35"
               24..72
               /gene="TCRB"
               join(24..72,161..449)
               /partial
               /note="putative"
               24..72
               /gene="TCRB"
               /note="T-cell receptor beta chain V-region, exon 1;
               G00-120-405; putative"
               <24..72
               /gene="TCRB"
               /note="T-cell receptor beta chain V-region; G00-120-405"
               /pseudo
               /number=1
               161..>449
               /note="T-cell receptor beta chain V-region"
               /number=2
               161..449
               exon

```

BASE COUNT 104 a 140 c 131 g 125 t
ORIGIN
/note="T-cell receptor beta chain V-region, exon 2; putative"
Query Match 0.2%; Score 24; DB 9; Length 500;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5855 GCAGAGAGGACAAAGCTGATGCT 5878
DB 158 GCAGAGAGGACAAAGCTGATGCT 135

RESULT 5
LOCUS HSDNATCRA/c 616 bp DNA linear PRI 26-JAN-1995
DEFINITION H.sapiens TCR V Beta 13.2 gene (allele a).
ACCESSION X75418
VERSION X75418.1 GI:472946
KEYWORDS TCR V-beta 13.2 gene.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 616)
AUTHORS Kay, R.A., Snowden, N., Hajeer, A.H., Boylston, A.W. and Ollier, W.E.
TITLE Genetic control of the human V beta 13.2 T cell repertoire:
Importance of allelic variation outside the coding regions of the
TCRBV13S2 gene
JOURNAL Eur. J. Immunol. 24 (11), 2863-2867 (1994)
MEDLINE 95045939
REFERENCE 2 (bases 1 to 616)
AUTHORS Kay, R.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-1993) R. Kay, University of Manchester, ARC
Epidemiology Research Unit, Stopford Bldg, Oxford Rd, Manchester
M13 9PT, UK

FEATURES
source
1. .616
/organism="Homo sapiens"
/isolate="patient 1"
/db_xref="taxon:9606"
/chromosome="7"
/map="q32-q35"
/cell_type="polymorph"
1. .616
/gene="TCR V Beta 13.2 (allele a)"
1. .231
/gene="TCR V Beta 13.2 (allele a)"
232. .281
/gene="TCR V Beta 13.2 (allele a)"
/number=1
233. .378
/gene="TCR V Beta 13.2 (allele a)"
/product="TCR V Beta 13.2"
join(233. .281,371. .>616)
/gene="TCR V Beta 13.2 (allele a)"
/codon_start=1
/product="TCR V Beta 13.2"
/protein_id="CAA53170.1"
/db_xref="GI:642237"
/translation="MSIGLLCGAFLSLMAGPVNAGVYQTQPKFRVLKTGSMTLCAQ
DMNHMYWYRQDPKGLRLIHYSVGEITAKGEVDPDGVNVSRLKQNFLLGLE"
282. .370
/gene="TCR V Beta 13.2 (allele a)"
/number=1
371. .>616
/gene="TCR V Beta 13.2 (allele a)"
/number=2
379. .>616
/gene="TCR V Beta 13.2 (allele a)"
/product="TCR V Beta 13.2"
mat_peptide

BASE COUNT 138 a 160 c 169 g 149 t
ORIGIN
Query Match 0.2%; Score 24; DB 9; Length 616;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5855 GCAGAGAGGACAAAGCTGATGCT 5878
DB 368 GCAGAGAGGACAAAGCTGATGCT 345

RESULT 6
LOCUS HSTRVB132/c 737 bp DNA linear PRI 05-APR-1992
DEFINITION H.sapiens TCR V-beta 13.2 gene for TCR beta chain variable region.
ACCESSION X61445
VERSION X61445.1 GI:37435
KEYWORDS T-cell receptor; TCR beta chain variable region; TCR V-beta 13.2
gene.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 737)
AUTHORS Li, Y., Szabo, P. and Posnett, D.N.
TITLE The genomic structure of human V beta 6 T cell antigen receptor
genes
JOURNAL J. Exp. Med. 174 (6), 1537-1547 (1991)
MEDLINE 92078864
REFERENCE 2 (bases 1 to 737)
AUTHORS Li, Y.
TITLE Direct Submission
JOURNAL Submitted (09-AUG-1991) Y. Li, Cornell Univ Medical College, 1300
York Avenue, Room D-601, Box 56, New York 10021, NY, USA

FEATURES
source
1. .737
/organism="Homo sapiens"
/isolate="human genomic library, Clontech labs., CAT#
HL1067J"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q35"
/clone_id="5-2"
/tissue_type="placental"
/clone_lib="human genomic library"
/germline
1. .266
267. .315
/gene="TCR V-beta 13.2"
/note="first leader"
/number=1
join(267. .315,404. .698)
/gene="TCR V-beta 13.2"
join(267. .315,404. .698)
/partial
/gene="TCR V-beta 13.2"
/codon_start=1
/protein_id="CAA43685.1"
/db_xref="GI:37436"
/translation="MSIGLLCGAFLSLMAGPVNAGVYQTQPKFRVLKTGSMTLCAQ
DMNHMYWYRQDPKGLRLIHYSVGEITAKGEVDPDGVNVSRLKQNFLLGSAP
SQTSVYFCASSTYS"
316. .403
/gene="TCR V-beta 13.2"
/number=1
404. .698
/gene="TCR V-beta 13.2"
/note="second leader and V region"
/number=2
170 a 195 c 189 g 183 t
BASE COUNT

ORIGIN

Query Match 0.2%; Score 24; DB 9; Length 737;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5855 GCAGAGAAGCAACGTCGTC 5878
|||||
DB 401 GCAGAGAAGCAACGTCGTC 378

RESULT 7
HSU07976/c 8045 bp DNA linear PRI 16-OCT-1994
LOCUS Human T cell receptor beta (TCRBV7S2, TCRBV13S2-1, TCRBV6S7-1)
DEFINITION genes, TCRBV deleted 2 haplotype, partial cds.
ACCESSION U07976
VERSION U07976.1 GI:558422
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 8045)
AUTHORS Zhao,T.M., Whitaker,S.E. and Robinson,M.A.
TITLE A genetically determined insertion/deletion related polymorphism in human T cell receptor beta chain (TCRB) includes functional variable gene segments
JOURNAL J. Exp. Med. 180, 1405-1414 (1994)
MEDLINE 95016432
REFERENCE 2 (bases 1 to 8045)
AUTHORS Robinson,M.A.
TITLE Direct Submission
SUBMITTED (24-MAR-1994) Mary Ann Robinson Ph.D., Laboratory of Immunogenetics, NIAID/National Institutes of Health, 12441 Parklawn Dr., Room 79, Rockville, MD 20852, USA
FEATURES
source Location/Qualifiers
1..8045
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q35"
/clone="D7S"
/haplotype="TCRBV deleted 2"
/sex="female"
/cell_line="SGMY"
/cell_type="B cell"
/gene="TCRBV7S2"
/join(311..359,469..>758)
/gene="TCRBV7S2"
/join(311..359,469..>758)
/gene="TCRBV7S2"
/codon_start=1
/product="T cell receptor beta"
/protein_id="AA50585.1"
/db_xref="GI:558423"
/translation="MGCRLCCAVLCIGAVPMETGYTORPHILVGMGNTKKSICQD HLGHNMYWKQSKAKKPLELMFYVNFKEQTEENSVPSRFSPECNSSHLPLHLTLDP EDSALVLCASS"
360..468
/gene="TCRBV7S2"
/join(4000..4048,4137..>4426)
/gene="TCRBV13S2-1"
/join(4000..4048,4137..>4426)
/gene="TCRBV13S2-1"
/codon_start=1
/product="T cell receptor beta"
/protein_id="AA50586.1"
/db_xref="GI:558424"
/translation="MSIGLCCGAFSLMAGPVNAGVYOTPKPFVYLKGTGOSMTLLCAQ DMNHEIYWIYRQDPGKGLRLIHVSVEGTTAKGVDPGVNYSRLKKONFLIGESAAP SQRSVFYWCASS"
4049..4136
/gene="TCRBV13S2-1"
Intron

gene join(7627..7675,7824..>8045)
/gene="TCRBV6S7-1"
/join(7627..7675,7824..>8045)
/gene="TCRBV6S7-1"
/codon_start=1
/product="T cell receptor beta"
/protein_id="AA50587.1"
/db_xref="GI:558425"
/translation="MGRPLLEWVAFCLIGADHPCAGVSQSPNSKVTKEGKQVELRCQD ISGHTALYWKQSKAKKPLELMFYVNFKEQTEENSVPSRFSPECNSSHLPLHLTLDP EDSALVLCASS"
7676..7823
/gene="TCRBV6S7-1"
Intron

BASE COUNT 2256 a 1813 c 1809 g 2167 t
ORIGIN

Query Match 0.2%; Score 24; DB 9; Length 8045;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5855 GCAGAGAAGCAACGTCGTC 5878
|||||
DB 4134 GCAGAGAAGCAACGTCGTC 4111

RESULT 8
HSU07978/c 33037 bp DNA linear PRI 16-OCT-1994
LOCUS Human T cell receptor beta (TCRBV7S2, TCRBV13S2-1a, TCRBV9S2p, TCRBV7S3, TCRBV13S2-1b, TCRBV6S7-1) genes, TCRBV inserted haplotype, partial cds.
DEFINITION genes, TCRBV inserted
ACCESSION U07978
VERSION U07978.1 GI:558429
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 33037)
AUTHORS Zhao,T.M., Whitaker,S.E. and Robinson,M.A.
TITLE A genetically determined insertion/deletion related polymorphism in human T cell receptor beta chain (TCRB) includes functional variable gene segments
JOURNAL J. Exp. Med. 180, 1405-1414 (1994)
MEDLINE 95016432
REFERENCE 2 (bases 1 to 33037)
AUTHORS Robinson,M.A.
TITLE Direct Submission
SUBMITTED (24-MAR-1994) Mary Ann Robinson Ph.D., Laboratory of Immunogenetics, NIAID/National Institutes of Health, 12441 Parklawn Dr., Room 79, Rockville, MD 20852, USA
FEATURES
source Location/Qualifiers
1..33037
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q35"
/clone="17s,17L,19,Ibg2,Ibg3,Ibg4"
/haplotype="TCRBV inserted"
/sex="male"
/cell_line="SBOR"
/cell_type="B cell"
/gene="TCRBV7S2"
/join(3734..3782,3892..>4181)
/gene="TCRBV7S2"
/join(3734..3782,3892..>4181)
/gene="TCRBV7S2"
/codon_start=1
/product="T cell receptor beta"
/protein_id="AA50590.1"
/db_xref="GI:558430"
/translation="MGCRLCCAVLCIGAVPMETGYTORPHILVGMGNTKKSICQD HLGHNMYWKQSKAKKPLELMFYVNFKEQTEENSVPSRFSPECNSSHLPLHLTLDP EDSALVLCASS"
3783..3891
Intron

```

gene      /gene="TCRBV7S2"
          join(7424..7472,7561..>7850)
          /gene="TCRBV13S2-1a"
          join(7424..7472,7561..>7850)
CDS       /gene="TCRBV13S2-1a"
          /codon_start=1
          /product="T cell receptor beta"
          /protein_id="AA50591.1"
          /db_xref="GI:558431"
          /translation="MSGLCCGAFSLMAGPVNAGVTQTPEFRVLTQGSMTLLCAQ
DNMHEMTWYRQDPGMGLRLIHVSVEGTAKGEVPDGVNVSRLKKNFLGLLESAP
SOTSVYFCASS"
          7473..7560
CDS       /gene="TCRBV13S2-1a"
          join(17328..17376,17494..>17783)
          /note="TCRBV9S2p"
          /codon_start=1
          /pseudo
          17377..17493
          /note="TCRBV9S2p"
          join(20774..20822,20932..>21221)
          /gene="TCRBV7S3"
          join(20774..20822,20932..>21221)
CDS       /gene="TCRBV7S3"
          /codon_start=1
          /product="T cell receptor beta"
          /protein_id="AA50592.1"
          /db_xref="GI:558432"
          /translation="MGCRLCCAVILCGAVPMETGYOTPRHLVGMTKKSLCKEO
HLSGNMVMYKOSAKKPLELMFYSLDERVNNVSFRFSPECNSSLHLHLHTLTP
EDSALYLCASS"
          20823..20931
          /gene="TCRBV7S3"
          join(28984..29032,29121..>29410)
          /gene="TCRBV13S2-1b"
          join(28984..29032,29121..>29410)
CDS       /gene="TCRBV13S2-1b"
          /codon_start=1
          /product="T cell receptor beta"
          /protein_id="AA50593.1"
          /db_xref="GI:558433"
          /translation="MSLGLCCGAFSLMAGPVNAGVTOTPKFRVLKTGQSMILLCAQ
DMHEMYWYRQDPGMGLRLIHVSVEGTAKGEVPDGVNVSRLKKNFLGLLESAP
SOTSVYFCASS"
          29033..29120
          /gene="TCRBV13S2-1b"
          join(32617..32665,32816..>33037)
          /gene="TCRBV6S7-1"
          join(32617..32665,32816..>33037)
CDS       /gene="TCRBV6S7-1"
          /codon_start=1
          /product="T cell receptor beta"
          /protein_id="AA50594.1"
          /db_xref="GI:558434"
          /translation="MGRILFWAVCLIGADHTGAGVSSPSNKYTKRKDVLEKCDP
ISGHTALWYRSLGGLEFLIFQGNAPSKGLSDFSAERIG"
          32666..32815
          /gene="TCRBV6S7-1"
          32666..32815
Intron    /gene="TCRBV6S7-1"
          32666..32815
BASE COUNT 9369 a 7118 c 7074 g 9476 t
ORIGIN

```

Query Match 0.2%: Score 24; DB 9; Length 33037;
 Best Local Similarity 100.0%; Pred. No. 2.1;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5855 GCAGAGAGACAAAGCTGATGCT 5878
 ||||||||||||||||||||||||||||
 Db 29118 GCAGAGAGACAAAGCTGATGCT 29095

RESULT 9
 AC101336
 LOCUS AC101336 72007 bp DNA linear HTG 23-NOV-2001

```

DEFINITION Mus musculus clone RP23-109p22, LOW-PASS SEQUENCE SAMPLING.
ACCESSION AC101336
VERSION AC101336.1 GI:17060111
KEYWORDS HTGS PHASED.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 72007)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Mus musculus, clone RP23-109p22
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 72007)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., All,A., Allen,N.,
Anderson,S., Barina,N., Bastien,V., Boguslavsky,L., Boukhalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cook,A.,
Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S., Fairo,S.,
Ferreira,P., Fitzhugh,M., Gage,D., Galagan,J., Gaidyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kellis,C., Larocque,K.,
Lamarez,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Melidim,J.,
Meneus,L., Mihova,T., Menga,V., Murphy,T., O'Donnell,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,J., O'Neill,D.,
Oliver,J., Peterson,R., Phunkhaph,P., Piere,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupbach,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Triggilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainou,R., Zembek,L., Zimmer,A. and Zody,M.
DIRECT SUBMISSION
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: 109_P_22
Center clone name: 109_P_22

```

* NOTE: This record contains 89 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

1 714: contig of 714 bp in length
 715 814: gap of 100 bp
 815 1492: contig of 678 bp in length
 1493 1592: gap of 100 bp
 1593 2314: contig of 722 bp in length
 2315 2414: gap of 100 bp
 2415 3136: contig of 722 bp in length
 3137 3236: gap of 100 bp
 3237 3958: contig of 722 bp in length
 3959 4058: gap of 100 bp
 4059 4794: contig of 736 bp in length
 4795 4894: gap of 100 bp

* 4895 5605: contig of 711 bp in length
* 5606 5705: gap of 100 bp
* 5706 6407: contig of 702 bp in length
* 6408 6507: gap of 100 bp
* 6508 7233: contig of 726 bp in length
* 7234 7333: gap of 100 bp
* 7334 8036: contig of 703 bp in length
* 8037 8136: gap of 100 bp
* 8137 8853: contig of 717 bp in length
* 8854 8953: gap of 100 bp
* 8954 9669: contig of 716 bp in length
* 9670 9769: gap of 100 bp
* 9770 10476: contig of 707 bp in length
* 10477 10576: gap of 100 bp
* 10577 11272: contig of 696 bp in length
* 11273 11372: gap of 100 bp
* 11373 12071: contig of 699 bp in length
* 12072 12171: gap of 100 bp
* 12172 12861: contig of 680 bp in length
* 12862 12961: gap of 100 bp
* 12962 13685: contig of 724 bp in length
* 13686 13785: gap of 100 bp
* 13786 14509: contig of 724 bp in length
* 14510 14609: gap of 100 bp
* 14610 15322: contig of 713 bp in length
* 15323 15422: gap of 100 bp
* 15423 16132: contig of 710 bp in length
* 16133 16232: gap of 100 bp
* 16233 16949: contig of 717 bp in length
* 16950 17049: gap of 100 bp
* 17050 17745: contig of 686 bp in length
* 17746 17845: gap of 100 bp
* 17846 18505: contig of 660 bp in length
* 18506 18605: gap of 100 bp
* 18606 19309: contig of 704 bp in length
* 19310 19409: gap of 100 bp
* 19410 20122: contig of 713 bp in length
* 20123 20222: gap of 100 bp
* 20223 20928: contig of 706 bp in length
* 20929 21028: gap of 100 bp
* 21029 21756: contig of 728 bp in length
* 21757 21856: gap of 100 bp
* 21857 22581: contig of 725 bp in length
* 22582 22681: gap of 100 bp
* 22682 23407: contig of 726 bp in length
* 23408 23507: gap of 100 bp
* 23508 24215: contig of 708 bp in length
* 24216 24315: gap of 100 bp
* 24316 25032: contig of 717 bp in length
* 25033 25132: gap of 100 bp
* 25133 25839: contig of 707 bp in length
* 25840 25939: gap of 100 bp
* 25940 26622: contig of 683 bp in length
* 26623 26722: gap of 100 bp
* 26723 27423: contig of 701 bp in length
* 27424 27523: gap of 100 bp
* 27524 28229: contig of 706 bp in length
* 28230 28329: gap of 100 bp
* 28330 29058: contig of 729 bp in length
* 29059 29158: gap of 100 bp
* 29159 29866: contig of 708 bp in length
* 29867 29966: gap of 100 bp
* 29967 30671: contig of 705 bp in length
* 30672 30771: gap of 100 bp
* 30772 31511: contig of 740 bp in length
* 31512 31611: gap of 100 bp
* 31612 32331: contig of 720 bp in length
* 32332 32431: gap of 100 bp
* 32432 33140: contig of 709 bp in length
* 33141 33240: gap of 100 bp
* 33241 33957: contig of 717 bp in length
* 33958 34057: gap of 100 bp
* 34058 34771: contig of 714 bp in length

* 34772 34871: gap of 100 bp
* 34872 35577: contig of 706 bp in length
* 35578 35677: gap of 100 bp
* 35678 36370: contig of 693 bp in length
* 36371 36470: gap of 100 bp
* 36471 37165: contig of 695 bp in length
* 37166 37265: gap of 100 bp
* 37266 37980: contig of 715 bp in length
* 37981 38080: gap of 100 bp
* 38081 38793: contig of 713 bp in length
* 38794 38893: gap of 100 bp
* 38894 39611: contig of 718 bp in length
* 39612 39711: gap of 100 bp
* 39712 40417: contig of 706 bp in length
* 40418 40517: gap of 100 bp
* 40518 41227: contig of 710 bp in length
* 41228 41327: gap of 100 bp
* 41328 42047: contig of 720 bp in length
* 42048 42147: gap of 100 bp
* 42148 42820: contig of 673 bp in length
* 42821 42920: gap of 100 bp
* 42921 43614: contig of 694 bp in length
* 43615 43714: gap of 100 bp
* 43715 44439: contig of 725 bp in length
* 44440 44539: gap of 100 bp
* 44540 45261: contig of 722 bp in length
* 45262 45361: gap of 100 bp
* 45362 46094: contig of 733 bp in length
* 46095 46194: gap of 100 bp
* 46195 46884: contig of 660 bp in length
* 46885 46984: gap of 100 bp
* 46985 47688: contig of 704 bp in length
* 47689 47788: gap of 100 bp
* 47789 48511: contig of 723 bp in length
* 48512 48611: gap of 100 bp
* 48612 49329: contig of 718 bp in length
* 49330 49429: gap of 100 bp
* 49430 50167: contig of 738 bp in length
* 50168 50267: gap of 100 bp
* 50268 50980: contig of 713 bp in length
* 50981 51080: gap of 100 bp
* 51081 51787: contig of 707 bp in length
* 51788 51887: gap of 100 bp
* 51888 52585: contig of 658 bp in length
* 52586 52685: gap of 100 bp
* 52686 53393: contig of 708 bp in length
* 53394 53493: gap of 100 bp
* 53494 54209: contig of 716 bp in length
* 54210 54309: gap of 100 bp
* 54310 55028: contig of 719 bp in length
* 55029 55128: gap of 100 bp
* 55129 55843: contig of 715 bp in length
* 55844 55943: gap of 100 bp

Query Match 0.2% Score 24; DB 2; Length 72007;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 5856 CAGAGAGACAAGCTGATGCTG 5879
|||||
Db 34396 CAGAGAGACAAGCTGATGCTG 34419

RESULT 10
AC022170
LOCUS
DEFINITION Homo sapiens chromosome 2 clone RP11-2001, WORKING DRAFT SEQUENCE,
3 unordered pieces.
ACCESSION AC022170
VERSION AC022170.14 GI:13122768
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULFILLTOP; HTGS_ACTIVEFIN.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 155021)
Abola, A.P., Bruno, D., Conn, L., Dela Rosa, M., Faulkner, D., Federspiel, N., Glukhov, S., Hansen, N., Herman, Z.S., Hyman, R., Mao, J., Lam, B., Marathe, R., Morehouse, A.J., Nguyen, M., Oefner, P., Palm, C.J., Ramirez, D., Southwick, A.M., Wilhelmy, J., Yu, S., and Davis, R.W.
Unpublished
2 (bases 1 to 155021)
Bruno, D., Conn, L., Dela Rosa, M., Federspiel, N., Foreman, P., Glukhov, S., Hansen, N., Hyman, R., Mao, J., Marathe, R., Morehouse, A.J., Oefner, P., Palm, C.J., Ramirez, D., Wilhelmy, J., Yu, S., and Davis, R.W.
Direct Submission
Submitted (26-JAN-2000) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
On Feb 24, 2001 this sequence version replaced gi:12863205.
----- Genome Center
Center: Stanford DNA Sequencing and Technology Development Center
Center code: SDSDC
Web site: <http://sequence-www.stanford.edu/group/human/>
Contact: hum-info@sequence.stanford.edu
----- Project Information
Center project name: 768
Center clone name: RP11-20G1
----- Summary Statistics
Sequencing Vector: M13mp18; X02513; 98% of reads
Sequencing Vector: plasmid; plasmid_accession: 2% of reads
Chemistry: Dye-terminator Big Dye; 92% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 153938 bases at least Q40
Consensus quality: 154052 bases at least Q20
Insert size: 11631; agarose-fp
Insert size: 154821; sum-of-contigs
Quality coverage: 12.3x in Q20 bases; agarose-fp
Quality coverage: 9.3x in Q20 bases; sum-of-contigs
* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as * runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
* 1 20875: contig of 20875 bp in length
* 20876 20975: gap of unknown length
* 20976 52837: contig of 31862 bp in length
* 52838 52937: gap of unknown length
* 52938 155021: contig of 102084 bp in length.
Location/Qualifiers
1. 155021
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/clone="RP11-20G1"
/clone_11p="RPC1 human BAC library 11"
1. 20875
/note="assembly_name:Contig33
clone_end:77"
misc_feature
20976..52837
/note="assembly_name:Contig34
clone_end:SP6"
52938..155021
/note="assembly_name:Contig35"
BASE COUNT 45493 a 33367 c 32844 g 43116 t 201 others
ORIGIN
Query Match 0.2%; Score 24; DB 2; Length 155021;

Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3390 ACACAAACATGACTGACGAAA 3413
DB 142072 ACACAAACATGACTGACGAAA 142095
RESULT 11
AC073415/c 215977 bp DNA linear HTG 24-JAN-2002
LOCUS Homo sapiens chromosome 2 clone RP11-84C2, WORKING DRAFT SEQUENCE,
DEFINITION 2 unordered pieces.
ACCESSION AC073415
VERSION AC073415.5 GI:18308794
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 215977)
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 215977)
Waterston, R.H.
Direct Submission
Submitted (16-JUN-2000) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
On Jan 24, 2002 this sequence version replaced gi:16195290.
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>
Contact: submissions@watson.wustl.edu
----- Project Information -----
Center project name: H_NH0084C02
----- Summary Statistics -----
Sequencing vector: M13; 42%
Sequencing vector: plasmid; 58%
Chemistry: Dye-terminator Big Dye; 42% of reads
Chemistry: Dye-terminator Big Dye; 58% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 214602 bases at least Q40
Consensus quality: 215177 bases at least Q30
Consensus quality: 215466 bases at least Q20
Insert size: 227000; agarose-fp
Insert size: 217039; sum-of-contigs
Quality coverage: 9.32 in Q20 bases; agarose-fp
Quality coverage: 8.26 in Q20 bases; sum-of-contigs
* NOTE: This is a 'working draft' sequence. It currently consists of 2 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as * runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
* 1 28725: contig of 28725 bp in length
* 28726 28825: gap of unknown length
* 28826 215977: contig of 187152 bp in length.
Location/Qualifiers
1. 215977
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/clone="RP11-84C2"
1. 28725
/note="assembly_name:Contig50
misc_feature
215977..28825
/note="assembly_name:Contig50"


```
misc.feature      clone_end:SP6
                  vector_side:right"
                  28826. .215977
                  /note="assembly_name:Contig51
                  clone_end:T7
                  vector_side:right"
BASE COUNT      61078 a 40739 c 43688 g 70372 t 100 others
ORIGIN
Query Match      0.24; Score 24; DB 2; Length 215977;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3390 ACACAAACATGACTGACAGAA 3413
Db 161277 ACACAAACATGACTGACAGAA 161254
|||||
|||||

RESULT 12
AX336388/c      267156 bp      DNA      linear      PAT 09-JAN-2002
LOCUS
DEFINITION      Sequence 6897 from Patent WO0194629.
ACCESSION      AX336388
VERSION
KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
REFERENCE
1 (bases 1 to 267156)
AUTHORS
Young,P.E., Augustus,M., Carter,K.C., Ebner,R., Endress,G.,
Horrigan,S., Soppet,D.R. and Weaver,Z.
Cancer gene determination and therapeutic screening using signature
gene sets
Patent: WO 0194629-A 6897 13-DEC-2001;
FEATURES
source
location/Qualifiers
1..267156
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT      76527 a 56343 c 55787 g 78499 t
ORIGIN
Query Match      0.28; Score 24; DB 6; Length 267156;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 5855 GCAGAGAAGCAGCAAGCTGCT 5878
Db 163714 GCAGAGAAGCAGCAAGCTGCT 163691
|||||
|||||

RESULT 13
U66059/c
LOCUS
DEFINITION      267156 bp      DNA      linear      PRI 30-NOV-1998
U66059.1 GI:1552494
Dopamine-beta-hydroxylase-like, TRY1, TRY2, TRY3, TCRBV27S1P,
TCRBV22S1A2N1T, TCRBV9S1A1T, TCRBV51A1T, TCRBV13S3,
TCRBV6S7P, TCRBV7S3A2T, TCRBV13S2A1T, TCRBV9S2A2PT, TCRBV7S2A1N4T,
TCRBV13S9/13S2A1T, TCRBV6S5A1N1, TCRBV50S1P, TCRBV13S1, TCRBV13S5,
TCRBV6S1A1N1, TCRBV13S2S1P, TCRBV5S5P, TCRBV13S1A1N1, TCRBV12S2A1T,
TCRBV21S1, TCRBV8S4P, TCRBV12S3, TCRBV21S3A2N2T, TCRBV8S5P,
TCRBV13S1 genes from bases 1 to 267156 (section 1 of 3).
U66059.1 36092
C region: C-beta gene segment; D region: J-beta gene segment;
J-segment: T-cell receptor beta chain; TCR-beta gene: V-beta gene
segment; V-segment: cell membrane protein; constant region;
diversity region: germline; joining segment; trypsin; trypsinogen;
variable segment.
Homo sapiens (clone: K41A) (clone_11b: CGM1: YAC D49H4) DNA; Homo
sapiens (clone: K35) (clone_11b: CGM1: YAC D49H4) DNA; Homo sapiens
(clone: K26) (clone_11b: CGM1: YAC D49H4) DNA; Homo sapiens (clone:
K56) (clone_11b: CGM1: YAC D49H4) DNA; Homo sapiens (library:
ATCC1475) (clone: K21B) (clone_11b: Kai Wang's) DNA; Homo sapiens
(library: ATCC1475) (clone: G54) (clone_11b: Kai Wang's) DNA; Homo
sapiens (library: Helai) (clone: H137) (clone_11b: Eric Lai's) DNA;
Homo sapiens (library: sperm) (clone: H18) (clone_11b: Eric Lai's)
DNA; Homo sapiens (clone: H18/G15gap) DNA; Homo sapiens (library:
ATCC1475) (clone: G15) (clone_11b: Kai Wang's) DNA; Homo sapiens
(library: ATCC1475) (clone: X1A) (clone_11b: Kai Wang's) DNA; Homo
sapiens (clone: A27) (clone_11b: CGM1: YAC 234 A6F6) DNA; Homo
sapiens (clone: A212partial) (clone_11b: CGM1: YAC 234 A6F6) DNA;
Homo sapiens (clone: A14) (clone_11b: CGM1: YAC 234 A6F6) DNA;
Homo sapiens (library: sperm) (clone: H7.1) (clone_11b: Eric Lai's)
DNA; Homo sapiens (clone: H12.18) (clone_11b: Eric Lai's) DNA;
Homo sapiens (library: Helai) (clone: H130) (clone_11b: Eric Lai's)
DNA; Homo sapiens (clone: A16) (clone_11b: CGM1: YAC 234 A6F6) DNA;
Homo sapiens (clone: C215) (clone_11b: CGM1: YAC 234 A7B3) DNA;
Homo sapiens (clone: G1) (library: ATCC1475) (clone_11b: Kai Wang's)
DNA; Homo sapiens (clone: C68) (clone_11b: CGM1: YAC 210 A38G1)
DNA; Homo sapiens (clone: C21) (clone_11b: CGM1: YAC 234 A7B3)
DNA; Homo sapiens (library: ATCC1475) (clone: X11) (clone_11b: Kai
Wang's) DNA; Homo sapiens (library: ATCC1475) (clone: X6A)
(clone_11b: Kai Wang's) DNA; Homo sapiens (library: ATCC1475)
(clone: CBG1) (clone_11b: Kai Wang's) DNA; Homo sapiens (clone:
CBG1/C29gap) DNA; and Homo sapiens (library: ATCC1475) (clone: C29)
(clone_11b: Kai Wang's) DNA.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
REFERENCE
1 (bases 1 to 267156)
AUTHORS
Rowen,L., Koop,B.F. and Hood,L.
The complete 685-kilobase DNA sequence of the human beta T cell
receptor locus
Science 272 (5269), 1755-1762 (1996)
JOURNAL
MEDLINE
66256474
2 (bases 1 to 267156)
AUTHORS
Rowen,L., Seto,J., Smit,A., Acharya,C., Ahearn,M.E., Ankeney,M.,
Baskin,D., Bumgarner,R., Chen,L., Chen,N., Deshpande,P., Faust,J.,
Howard,S., Jerome,N., Koop,B.F., Lee,H., Loretz,C., Paeger,B.,
Zackroff,K. and Hood,L.
Sequence determination of the human T cell receptor beta locus:
Strategy and error analysis
Unpublished
JOURNAL
3 (bases 1 to 267156)
AUTHORS
Rowen,L., Wang,K., Boyse,C., Ahearn,M.E., Charmley,P., Paeger,B.,
Lee,I., Chen,L., Trask,B., Nickerson,D., Seto,D. and Hood,L.
Sequence variation among several haplotypes in the human T cell
receptor beta locus
Unpublished
JOURNAL
4 (bases 1 to 267156)
AUTHORS
Rowen,L.
Direct Submission
Submitted (15-OCT-1994) L.Rowen leetrowen@u.washington.edu
5 (bases 1 to 267156)
AUTHORS
Rowen,L.
Direct Submission
Submitted (24-JUN-1997)
This sequence overlaps section 2 of the T cell receptor beta locus,
Genbank Accession Number U66060, by 9647 bases.
FEATURES
source
location/Qualifiers
1..267156
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="7q35"
/complement(1..240)
/note="putative"
/rpt_family="Alu"
1..37031
/note="(vector PWE15A)"
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="K41A"
/cell_line="CGM1, haplotype B"
/clone_11b="YAC D49H4"
repeat_region
source
```

```

repeat_region /germline
complement(577..709)
/nc="SINE; putative"
repeat_region /rpt_family="MIR"
complement(1463..1575)
/nc="MALR retroposon LTR; putative"
repeat_region /rpt_family="MTIG"
2307..2428
/nc="DNA transposon fossil; putative"
repeat_region /rpt_family="MER33; fragment 1"
2429..7244
/nc="LINE; putative"
repeat_region /rpt_family="L1PA7; fragment 1"
2429..2988
/nc="putative"
repeat_region /rpt_family="LINE 1"
7245..7975
/nc="LINE; putative"
repeat_region /rpt_family="L1PA7; fragment 2"
7997..8294
/nc="putative"
repeat_region /rpt_family="Alu"
8295..8463
/nc="putative"
repeat_region /rpt_family="MER33"
8764..8793
/nc="15 bp tandem duplication; putative"
repeat_region /rpt_type=tandem
complement(9085..9224)
/nc="DNA transposon fossil; putative"
repeat_region /rpt_family="MER33"
9888..9965
/nc="SINE; putative"
repeat_region /rpt_family="MIR"
10076..10320
/nc="LINE; putative"
repeat_region /rpt_family="L1MA4; fragment 1"
10321..10616
/nc="putative"
repeat_region /rpt_family="Alu"
10617..10747
/nc="LINE; putative"
repeat_region /rpt_family="L1MA4; fragment 2"
12541..13965
/nc="60 (120) mer, 19x repeated, 91% conserved first
5.5 120 mers; putative"
repeat_region /rpt_type=tandem
15873..16381
/nc="DNA transposon fossil; putative"
repeat_region /rpt_family="MER64"
complement(19303..19385)
/nc="SINE; putative"
repeat_region /rpt_family="MIR"
19342..19632
/nc="SINE; putative"
repeat_region /rpt_family="MIR"
19682..19755
/nc="LINE; putative"
repeat_region /rpt_family="L1MA3"
19945..20437
/nc="retroposon LTR; putative"
repeat_region /rpt_family="MER65B"
complement(21066..21541)
/nc="retroposon LTR; putative"
repeat_region /rpt_family="MER60A"
complement(22822..23202)
/nc="LINE; putative"
repeat_region /rpt_family="L1PA16"
23444..23601
/nc="SINE; putative"
repeat_region /rpt_family="MIR"
complement(23615..23909)
/nc="putative"

```

```

repeat_region /rpt_family="Alu"
2403..24079
/nc="SINE; putative"
repeat_region /rpt_family="MIR"
complement(24304..24697)
/nc="endogenous retroviral LTR; putative"
repeat_region /rpt_family="LTR1; fragment 1"
complement(25086..25502)
/nc="endogenous retroviral LTR; putative"
repeat_region /rpt_family="LTR1; fragment 2"
complement(25975..26233)
/nc="putative"
repeat_region /rpt_family="Alu"
27165..28630
/nc="LINE; putative"
repeat_region /rpt_family="L1P5; fragment 1"
28056..28507
/nc="putative"
repeat_region /rpt_family="LINE 1"
28631..28702
/nc="LINE; putative"
repeat_region /rpt_family="L1P5; fragment 2"
complement(28713..29101)
/nc="LINE; putative"
repeat_region /rpt_family="MIR2"
complement(30772..31746)
/nc="LINE; putative"
repeat_region /rpt_family="MER42C"
complement(31996..32116)
/nc="DNA transposon fossil; putative"
repeat_region /rpt_family="MER5B"
32313..32419
/nc="LINE; putative"
repeat_region /rpt_family="MER42C"
complement(33276..33406)

```

Query Match 0.2% Score 24; DB 9; Length 267156;
Best Local Similarity 100.0%; Pred. NO. 1.8;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5855 GCAGAGAGGACAAAGCTGATGCT 5878
DB 163714 GCAGAGAGGACAAAGCTGATGCT 163691
|||||

RESULT 14
LOCUS AP001119 293181 bp DNA linear BCT 07-SPP-2000
DEFINITION Buchnera sp. APS genomic DNA, complete sequence, segment 2/2.
ACCESSION AP001119 AP000398 BA000003
VERSION AP001119.1 GI:10038987
KEYWORDS
SOURCE Buchnera sp. APS (strain:APS isolate:Tokyo1998,
specific_host:Acyrthosiphon pisum) DNA.
ORGANISM Buchnera sp. APS
Bacteria; Proteobacteria; gamma subdivision; Buchnera.
REFERENCE 1 (sites)
AUTHORS Shigenobu,S., Watanabe,H., Hattori,M., Sakaki,Y. and Ishikawa,H.
TITLE Genome sequence of the endocellular bacterial symbiont of aphids
Buchnera sp. APS
JOURNAL Nature 407 (6800), 81-86 (2000)
MEDLINE 20445173
REFERENCE 2 (bases 1 to 293181)
AUTHORS Shigenobu,S., Hattori,M., Watanabe,H., Yada,T.,
Sakaki,Y. and Ishikawa,H.
TITLE Submitted (24-JAN-2000) Shuji Shigenobu, University of Tokyo,
Department of Biological Sciences; 7-3-1 Hongo, Bunkyo-ku, Tokyo
113-0033, Japan (E-mail:shige@sc.riken.go.jp,
URL:http://buchnera.gsc.riken.go.jp, Tel:81-3-5800-3553,
Fax:81-3-5800-3553)
COMMENT Collaboration Information:
RIKEN, Genomic Sciences Center (GSC)

Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
Japan
Tel:81-42-778-9923
FAX:81-42-778-9924
AP000398: Submitted(23-Aug-1999).

FEATURES

source

Location/Qualifiers

1..293181

/organism="Buchnera sp. APS"

/strain="APS"

/isolate="Tokyo1998"

/specific_host="Acyrthosiphon pisum"

/db_xref="taxon:107806"

/note="primary endosymbiont of Acyrthosiphon pisum"

321..1280

/gene="crxb"

321..1280

/gene="crxb"

/EC_number="1.6.4.5"

/note="BU314"

/codon_start=1

/transl_table=11

/product="thioredoxin reductase"

/protein_id="BAB13022.1"

/db_xref="GI:10038988"

/translation="MDKVRHKKIILIGSPACATTAIYARANDLPFLITGKNGCOL
MNTNIEWMPGDYKISGSELNNRMKHAIEKTYVICDPTVISVNFKNPFLLIGBN
KYTADSVIATGPNRYLGLQSELEFKGAVSTCAVCDGFEYKNKEVAAGGNTAIE
ETLYSNVVKKHLHRCINFRAEKILDLRLKIKSOKIITYLNSIYVNIAGNSGV
TALLIEOKNSKEKTESKIOVSGLFVAISGTPMNTNFVNLKKDGIQVTRQEHGNYT
OTSIPIGAAGVIDHYVRQAITSSASGCMALDSERYINSLV"

1387..1605

/gene="infa"

1387..1605

/gene="infa"

/note="BU315"

/codon_start=1

/transl_table=11

/product="translation initiation factor IF-1"

/protein_id="BAB13023.1"

/db_xref="GI:10038989"

/translation="MSKEENIMOGVVIDTLPTMFRVLEKNHITYTAHSGMKRKY
IRLLGDKVTVELTPYDLTKGRITFRSR"

1774..3534

/gene="asps"

1774..3534

/gene="asps"

/EC_number="6.1.1.12"

/note="BU316"

/codon_start=1

/transl_table=11

/product="aspartyl-tRNA synthetase"

/protein_id="BAB13024.1"

/db_xref="GI:10038990"

/translation="MRTKCGNIRISHNNKKKYLKCGVHKNVNLGQFIVDKDRTGL
VQVIELKNYKIEFKALNIRNEFCIQVEFTVOKREKKNNIKIRTGEIEILANVLIT
NTSKPLNFTGENDNDRLKRYLDFLRSFDILENLKIRNKITYLIRNMTKKNFEDI
ETPLITKSTPEGARDYLVPSRNHYGKFLAPQSPOLKFOILMISGIDRRYQVCKEFDI
EDLRSDRPEFTQIDIEVSEFSAKKIRNLVENLTKIKLMEIRINILKKRPOISPHAM
KRGSDKRDRLRNPIETIDVSNTEFKDKKTFSPNLNOKNRITALLICISGAHLSRKKI
DDYTRYVORFDKAKLFYMKIKCKLGLGSHSITNIDELITLKEITKESQSKNDIL
FLIAQDEHIVNSLGMRLKIGIDILNITKKNWEPDWLVNPFMDKDIOGNISVHPH
ETVIANMDREILKNSPDLAISDSYDLINGEYIGGVSRIHDVNMQKOVFDIIGIKKS
MONEKFGFLIEALKYCAPPHAGIALGLDRIVMLITNSKNIRDVIAFPKTTATCLMTN
SPSYVDNLLOLAIKHLKK"

complement(3545..4333)

/gene="znub"

complement(3545..4333)

/gene="znub"

/note="BU317"

/codon_start=1

/transl_table=11

/product="high-affinity zinc uptake system membrane
protein znub"

gene
CDS

complement(3545..4333)
/gene="znub"
complement(3545..4333)
/gene="znub"
/note="BU317"
/codon_start=1
/transl_table=11
/product="high-affinity zinc uptake system membrane
protein znub"

/protein_id="BAB13025.1"
/db_xref="GI:10038991"
/translation="MPELIFPGWLACVLLSLTTGPGCSFTVWRMSSFEDTULSHSLL
GIALSIAFNINSFYAILIMSFATILAMLEELPVLDTVLNIIISHSSLSGMVTS
LISSKKEINTITNYLFGDLISVFNKNDLITTSISLITLSTLLFPWHSILSTINEELSO
INGINLVARITLMTLAFITAIKAEVAGLITSLTLLIPPAOAHFSSGPEKMWITA
IIVSILSVTGGISLSVFYMPSPASPSIVLCSSFCLSLNKKHRY"
complement(4391..5107)
/gene="znuc"

complement(4391..5107)

/gene="znuc"

/note="BU318"

/codon_start=1

/transl_table=11

/product="high-affinity zinc uptake system ATP-binding
protein znuc"

/protein_id="BAB13026.1"

/db_xref="GI:10038992"

/translation="MPEFYKLRKNCVNLSHRSITLPDISLSISNRVLTLLGPNCAKS
TLVRITLGLKPNSGTITIRSNLSVGYVPOKHLNLTLLPTVBRFKLSRKTNNITE
EMLRKVASLRFQQLQKLSGEMORILFAKALNPNLLVDEPTQGVNGQALY
KLINERHELOCAILIVSHDLNFMVAKTDVLCINNHICSGTPEVCNNLFEISIFG
LKVQELATYHHNNHHNF"

complement(6425..7867)

/gene="pyka"

complement(6425..7867)

/gene="pyka"

/EC_number="2.7.1.40"

/note="BU319"

/codon_start=1

/transl_table=11

/product="pyruvate kinase"

/protein_id="BAB13027.1"

/db_xref="GI:10038993"

/translation="MLNRLRRKIVATLGPSTDINNLEKIIRSGVNLRFNFSHGS
DEKSRANORAEIMPLRNFIHALDGLQPKIRSEIKNNIFLNGDFFITLDANDQ
NNGNDRVGSIDYKOLPYDLKGVGVDLLDGRIDQKIKSTGHEITLKYIGTILSNK
GINKLGGISADALEKDKRDIIILATEINVDYLAISFPSCSNDLQARKLAEFGSNA
KIITAKTERAEAVLNONTIEDIILASDAIMARODLGEVIGDSLQIKRLRTAROL
NRIVITATOMESMITNPLPTRAEVDVANNAYLDGSDVAMLSETASGEYPAETVKM
AKICKAEKVPISNVSRRHAKFDDIEEALMSAVYVANHUKGITAITTMTESGKTA
LMSRTSGLPFLALSKNKEETLNLASLYGVPYPIYNSKNNNKANENAVILRRKGF
LCRGLVITITGDMIGKIGKTINTSRILKVV"

8150..9625

/gene="zwf"

8150..9625

/gene="zwf"

/EC_number="1.1.1.49"

/note="BU320"

/codon_start=1

/transl_table=11

/product="glucose-6-phosphate 1-dehydrogenase"

/protein_id="BAB13028.1"

/db_xref="GI:10038994"

/translation="MIEETNHACDLVIFGAKGDLTKRKLLPALYKLESKKTHKYTRI
IASGRDAMTDEYEIEKIKTEVKNFLNEEINDILNKLRSRIFPCNDIVPEPLFRKL
TILKOKKNIIYVCAVPSNTLSIFIGLGNHLNLSVPSRIVLEKPLGVCLTKSKIND
QISXYLESDQIFRIDHYLGKESILNLEFALRFSTYCLFYMNKNTIDHIIQYSENGI
DRWNYFNMGOMKMDVONHLLQITLIMDQPKNTSSFSIOHEKKYILRSINPNIH
NINKKTVRQCYGCVYNEKKVPSYLEENAKNNSCTETVVAIKVDLNNKQSGVPYTL
RIGKRLTAHRYSEIVVEFKRPTNLFNLSLQNNKLLIRLEPNPTITDFSVKAGL
EODEYKILNYSLSQFSKRYSKNSIDAYERLLFEIMRGVQSLFVCHDELTAAMKWIDP
IHAMNSKNNAPOLYMSGTGWPKNSDLLLADHGRAWYEFH"

9810..10688

/gene="htpx"

9810..10688

/gene="htpx"

/note="BU321"

/codon_start=1

/transl_table=11

/product="heat shock protein htpx"

/protein_id="BAB13029.1"

/db_xref="GI:10038995"

REFERENCE	MO 63108, USA
AUTHORS	University School of Medicine, 4444 Forest Park Parkway, St. Louis
TITLE	6 (bases 1 to 58777)
JOURNAL	Waterston, R.
REFERENCE	Direct Submission
AUTHORS	Submitted (28-MAR-2000) Department of Genetics, Washington
TITLE	University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
JOURNAL	7 (bases 1 to 58777)
REFERENCE	Waterston, R.
AUTHORS	Direct Submission
TITLE	Submitted (19-JUL-2001) Department of Genetics, Washington
JOURNAL	University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
REFERENCE	Louis, MO 63110, USA
AUTHORS	8 (bases 1 to 58777)
TITLE	Waterston, R.
JOURNAL	Direct Submission
REFERENCE	Submitted (03-OCT-2001) Department of Genetics, Washington
AUTHORS	University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
TITLE	Louis, MO 63110, USA
JOURNAL	9 (bases 1 to 58777)
REFERENCE	Waterston, R.
AUTHORS	Direct Submission
TITLE	Submitted (07-DEC-2001) Department of Genetics, Washington
JOURNAL	University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
REFERENCE	Louis, MO 63110, USA
AUTHORS	On Mar 1, 2000 this sequence version replaced gi:4263205.
TITLE	Submitted by:
JOURNAL	

QY 3957 AAACATTGAGCAGTCAATTGAAA 3980

|||||
AAACATGAGCAGTCAATTGAA

AC006733

Accession	Length	Type	Release Date
AC006773	58777 bp	DNA linear	INV 07-DEC-2001
Caenorhabditis elegans cosmid Y32H12a, complete sequence.			
AC006773			

AC006733.2 GI:7140307
HTG.

ORGANISM

Eukaryota; Metazoa; Chromadorea; Rhabditida; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis (bases 1 to 5977).

1 (bases 1 to 58777)
The *C. elegans* Sequencing Consortium.
Genome sequence of the nematode *C. elegans*: a reference genome.

genome sequence of the nematode *C. elegans*: a platform for investigating biology. The *C. elegans* Sequencing Consortium Science 282 (5396), 2012-2018 (1998)

2 (bases 1 to 58777)

Holmes, A., Elliot, G. and Cloude, J.
The sequence of *C. elegans* cosmid Y32H12A
nucleotide

Unpublished
3 (bases 1 to 58777)
Waterston P

waterston, R.
Direct Submission
Unpublished

unpublished
4 (bases 1 to 58777)
Waterston, R.H.

Submitted (23-FEB-1999) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St.
MO 63108, USA

5 (bases 1 to 58777)
Waterston, R.H.
Direct Subjection

Direct Submission
Submitted (01-MAR-2000) Genome Sequencing Center, Washington

Submitted by:

Genome Sequencing Center
Department of Genetics, Washington University
St. Louis, MO 63110, USA, and
Sanger Centre, Hinxton Hall
Cambridge CB10 1RQ, England
Email: twenemacode.wustl.edu and jes@sanger.ac.uk

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one m13 subclone.

Euk

This is a segment of YAC Y22H12 that was sequenced to span the gap between ZK121 and W03A5. The 5' cosmid is CEL2121, 800 base overlap, the 3' is clone CELW03A5, 200 base overlap. Actual start of this clone is at base position 6240 of CELZK121; actual end is at 32064 of CELW03A5.

220

Coding sequences below are predicted from computer analysis, using the program Genefinder (P. Green and L. Hillier, ms in preparation)

3

Wat

Subr

MO
CH

13794

Dire

```

/odon_start=1
/product="Hypothetical protein Y32H12A.2"
/protein_id="AAF60487.1"
/db_xref="GI:7331799"
/translation="MSVNVPRPKRVREEDQKKRNPEIKISEKATQLTITDMPKMSDEKLT
FMLEDKATMGSDRIFFEEFLQDYAFLRNAVEEYLDGKPTSSDTISKIARLRNRTLT
HGTATLMRKREVAEMENDVAKLLQONVTSVOHIOKEIDICDLDSFAGSEELVEE
OPHTASSELANETADNEHOYLARKPENBORRELTSLNLEGRNAGELODIRKE
VRLQIGKREKLEDLTLYAEPVEFMGAKFKDLSTEGEORKLSTQLPAPLVAHITHAAY
KEIHGEYSKILRNFLYKYLEFEFLGINDTEKOSKDETEAMKRRKRDDETTQEKIN
KMASELLEVHPMSLEFEIECSSENIKISMQIQLTELKVTAKMHIWTDSEKLSKSPFS
IDLSMSEFENDTGDKCPMDVGEMKEIHKMNSNSAKOIGKPFKIOISSECTD
RIGATEHRLRKVTITTRTRISNVAEMDGIIROLESKRTDNLLEPEVKINSFICDES
FIASITPTSTIKISNSAADRSYKTEASDNGRKLTATISIPADYPRHTALEGIAP
EDSESSTEARQIEDRLNTEAHYVDVKSIRKQSLSLKHHIDS"
/length=15954
/genes="Y32H12A.1"
/join(15954..16288,16538..16743,17240..17524,17806..17912)
/genes="Y32H12A.1"
/notes="similar to C. elegans family of proteins; see
(GI:281088)"
/codon_start=1
/product="Hypothetical protein Y32H12A.1"
/protein_id="AAF60488.1"
/db_xref="GI:7331800"
/translation="MSGATYICCRVFNIRIKIVKMSYIYSKMSSEHESANATFVQCP
PTEHOILAIVLMSIKILNSIGLNLCLMLWILISYQCLNHTFSILTYTLINDCLI
SVTLTLTKRIVQCHETLAMEFGPCTFSTSEFCVLDAHLHALQNLWLSQILIGF
RTAMLSASQSFISMKRTITFESEYSPIDPFTAMLLINPIHPIFLAYCFRHHON
NDLKRRTMSMTLRMHSSTLRKYSTKAKFLFLDLVLVLPFCQLELPPQFLEVFLG
TISSINILNPYAMCCFRPYRKMEKREYRLIIS"
/complement(18325..21785)
/genes="Y32H12A.3"
/complement(join(18325..18714,20353..20604,20727..20992,
21734..21785))
/genes="Y32H12A.3"
/notes="contains similarity to pfam family PF00106 (short
chain dehydrogenase), score=124.1, E=2.6e-33, N=1; coded
for by the following C. elegans cDNAs: yk98a9.3, yk98a9.5,
yk216d4.3, yk216d4.5, yk229h3.3, yk229h3.5, yk274b4.3,
yk274b4.5, yk290g10.3, yk290g10.5, yk325b5.3, yk325b5.5,
yk325d6.3, yk325d6.5, yk374b9.3, yk428d1.3, yk428d1.5,
yk460f3.5, yk468b12.3, yk468b12.5"
/codon_start=1
/product="Hypothetical protein Y32H12A.3"
/protein_id="AAF60486.1"
/db_xref="GI:7331798"
/translation="MSLAGQIIVTASGSRIGRGIALOLGEAGATVYITGRKPEESLN
SKVLSGLEATADDETTRKGGCIARFVDHQNMEPVKNEPVEYKEKHQGLDILVNAY
OGVTAISEMCKPFYETDPIYWDITNNNGLRNHHICTYAAALMTARNKGLIVNYSG
GGLRLFLNVAVGVQALDRMSADTAVELRRKNVCVVSIMPGAVRELVDKMFKDENG
KRPRIKNAVEFANGETVEYPGRAVVSILASDRMDKGRILITDELKEYGXFVDIDG
LRPPNLRVSSEFLIKHGWNTANFVPTWKLPGMLVMAGNSRL"
/complement(21798..21799)
/notes="SUI trans-splice site; see EST yk98a9.5"
/genes="Y32H12A.4"
/complement(join(22928..23067,23835..23901,23954..24106,
24795..25016))
/genes="Y32H12A.4"
/notes="contains similarity to human protein phosphatase
inhibitor (SP:P41236); coded for by the following C.
elegans cDNAs: yk289g2.3, yk289g2.5, yk464f10.3,
yk464f10.5"
/codon_start=1
/product="Hypothetical protein Y32H12A.4"
/protein_id="AAF60485.2"
/db_xref="GI:14916388"
/translation="MASPSASADVDYKLTADGKICLDPAEPLKRRPKKSILMKODSS
LEGKDGRAHFDEMNILATYHPTDKDYSKMSIDEKPTYNHSDGESECDGLGVPTRR
PRRVSLGNALDEPKVAEGLAHPESGSSASDSDSEADLTTEORAHNRDPEFKRRRAH
YNEGAAALKHHPLEDEDEEATAAAGAGNMMEH"
/complement(29721..33664)

```

```

CDS
/genes="Y32H12A.5"
/complement(join(29721..29909,29971..30183,30397..31131,
31540..31907,32590..32729,32778..32881,33414..33503,
33581..33664))
/genes="Y32H12A.5"
/notes="contains similarity to schizosaccharomyces pombe
hypothetical protein (SP:I009749); coded for by the
following C. elegans cDNAs: cm17a5, yk222c2.5, yk345c9.5,
yk356h10.5"
/codon_start=1
/product="Hypothetical protein Y32H12A.5"
/protein_id="AAF60489.1"
/db_xref="GI:7331801"
/translation="MEEDVESSTPAESQKLLQSVNSFDERDMVVLNDELIMEN
RNSNSILPERSVTVDVLRARENITTDGPIQLADTQPNQVYIDQVVOFPNSQPSQO
REETIVRHSKVVVEQPASQQRDESMVYTRQNRPLETLEGLEIDGDPVSTSKAALPFK
KGRRAMSPNNAKGEKMTVAVVIDHNEEGGNORRRIYVYKLLHPQRRGHEDKHLRFV
TAGHDAEIHLDSDDDDELEVEINEEVIIPSTGEGPRAVIRKRFEMAKMTNFETL
PEMLQDNEYLRTHGHRPRLPFSQCFKSFALHETGNIHMVGCVAFFGIGTWPLTR
PHQOMQWVEKLIYSAFLCAVTCGSEFLFHHYACHISITMAKLFSLDTGTGSLIYVC
SFVWITFAFVCPQPMIITNGMIIITLGIAMVSLMDXFAEPKFRVRAAVFIAMGL

```

Search completed: October 7, 2002, 17:59:30
Job time : 17827 secs

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 7, 2002, 08:58:48 ; Search time 218.88 Seconds
(without alignments)
10802.577 Million cell updates/sec

Title: US-09-150-867-2

Perfect score: 9626
Sequence: 1 gaattccggagtcgagatag.....ttaaaaaaaaaacggaattc 9626

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5B.COMB.seq: *
2: /cgn2_6/ptodata/2/ina/6A.COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq: *
5: /cgn2_6/ptodata/2/ina/6B.COMB.seq: *
6: /cgn2_6/ptodata/2/ina/6B.COMB.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	95.8	1.0	3741	4	US-09-541-782-9	Sequence 9, Appl
2	90	0.9	29793	4	US-09-302-812-38	Sequence 38, Appl
3	90	0.9	29793	4	US-09-511-477-38	Sequence 38, Appl
4	88	0.9	5093	1	US-08-468-036-23	Sequence 23, Appl
5	88	0.9	5093	2	US-08-376-843-23	Sequence 23, Appl
6	85.2	0.9	7218	1	US-08-232-463-14	Sequence 14, Appl
7	82.8	0.9	1077	4	US-09-722-139-3	Sequence 3, Appl
8	82.8	0.9	4176	4	US-09-722-139-1	Sequence 1, Appl
9	82	0.9	1230	4	US-09-572-191-5	Sequence 5, Appl
10	82	0.9	4757	4	US-09-572-191-1	Sequence 1, Appl
11	79.8	0.8	2389	1	US-08-480-552-18	Sequence 18, Appl
12	79.8	0.8	2389	2	US-08-486-382-4	Sequence 4, Appl
13	79.8	0.8	2389	3	US-09-235-546-4	Sequence 4, Appl
14	79.8	0.8	2389	4	US-08-929-208-18	Sequence 18, Appl
15	79.8	0.8	2389	3	US-09-568-315-18	Sequence 18, Appl
16	78	0.8	1128	4	US-09-572-191-3	Sequence 3, Appl
17	73.8	0.8	3572	2	US-08-713-815A-2	Sequence 2, Appl
18	72.8	0.8	2870	1	US-08-468-036-28	Sequence 28, Appl
19	72.8	0.8	2870	2	US-08-376-843-28	Sequence 28, Appl
20	68	0.7	5361	4	US-08-973-462-2	Sequence 2, Appl
21	68	0.7	6152	4	US-08-973-462-1	Sequence 1, Appl
22	67.6	0.7	3930	3	US-09-162-373-2	Sequence 2, Appl
23	67.6	0.7	3930	4	US-09-467-946-2	Sequence 2, Appl
24	63.4	0.7	2387	4	US-09-723-153-1	Sequence 1, Appl
25	62.2	0.6	3883	1	US-08-468-036-33	Sequence 33, Appl
26	62.2	0.6	3883	2	US-08-376-843-33	Sequence 33, Appl
27	62.2	0.6	3884	4	US-09-541-782-3	Sequence 3, Appl

28	59.8	0.6	3095	6	US-09-541-782-1	Patent No. 5231168
29	58.6	0.6	3709	4	US-09-328-111-442	Sequence 7, Appl
30	58.4	0.6	1002	4	US-09-328-111-442	Sequence 442, App
31	56	0.6	1014	4	US-09-723-153-3	Sequence 3, Appl
32	55.8	0.6	897	4	US-09-621-233-1	Sequence 1, Appl
33	55	0.6	3489	2	US-08-728-323A-1	Sequence 1, Appl
34	55	0.6	5433	3	US-08-929-329-1	Sequence 1, Appl
35	55	0.6	32207	2	US-08-770-379-20	Sequence 20, Appl
36	55	0.6	32207	4	US-08-757-669A-20	Sequence 20, Appl
37	55	0.6	32207	4	US-09-230-371A-20	Sequence 20, Appl
38	54	0.6	2277	1	US-08-676-967-2	Sequence 2, Appl
39	54	0.6	2277	2	US-08-676-974-2	Sequence 2, Appl
40	54	0.6	2277	2	US-09-098-487-2	Sequence 2, Appl
41	52.6	0.5	4014	4	US-09-541-782-1	Sequence 1, Appl
42	50.2	0.5	2223	1	US-08-257-073-4	Sequence 4, Appl
43	48	0.5	2188	1	US-07-865-662E-10	Sequence 10, Appl
44	48	0.5	2188	4	US-08-374-219B-10	Sequence 10, Appl
45	47.2	0.5	1070	4	US-09-718-841-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1	US-09-541-782-9
Sequence 9, Application US/09541782	
Patent No. 6284480	
GENERAL INFORMATION:	
APPLICANT: Nislow, Corey	
APPLICANT: Sakowicz, Roman	
APPLICANT: Beraud, Christophe	
TITLE OF INVENTION: Antifungal Assay	
FILE REFERENCE: 1015	
CURRENT APPLICATION NUMBER: US/09/541, 782	
CURRENT FILING DATE: 2000-04-03	
NUMBER OF SEQ ID NOS: 10	
SOFTWARE: FastSeq for Windows Version 4.0	
SEQ ID NO 9	
LENGTH: 3741	
TYPE: DNA	
ORGANISM: H. sapiens	
US-09-541-782-9	
Query Match	1.0%; Score 95.8; DB 4; Length 3741;
Best Local Similarity	53.5%; Pred. No. 6.8e-14;
Matches 286; Conservative	0; Mismatches 222; Indels 27; Gaps 3;
QY 684 aaaaaggtgtaaaaaaacagcattatgagagactaaatgaatgacatagctagctgct	
DB 654 aaaaaggtgtaaaaaaacagcattatgagagactaaatgaatgacatagctagctgct	
QY 744 caatacaaatattgaatgatgtgtgaaagccgagacagaaatgacccaatctag	
DB 714 ccaactcagcttctcctgtaacaatacatatgaaga-----aaactcga	
QY 804 agaacgtgtagagcgtcgtcatgtatctcaactgaaattgtgtagatctgcggcagtg	
DB 759 ttgatgtgagaagcgtgtgtaaaatcggaaagtgtgaactgtgtgtcgtcggagtg	
QY 864 aagaagcaaaccaaacggtggaaggtgtggaacttaaggaaggtcgaagaaggtcgaagaac	
DB 819 aaaaacttgccgtctcggagcgtgtgtaagaagagctcgggaagctcgaaataataac	
QY 924 gcaagctgttatcccttgagcaggttataagaagacttaagcagcggccagcgtgagat	
DB 879 aatccctgtgactttgggaaggttcattctcctcttgagaagaacccctcag---	
QY 984 tttaaacatcagagacagcaaacctcagcaagaattctccaataatcattctgaggaag	
DB 936 ---tcttaccgagatcctaaactaactagaaatcctccagatctctcttgagggcgta	
QY 1044 ctaaaacgttataatttgcaacaattacgcaatttc-----tttgatgagaacctaa	

```

Db 993 caagacacatataattgcaacaatttctctgcctctcaattcttgaggaacctcga 1052
Qy 1098 gfaccttcagttgcccagctactcccaacatgtgagaataactcctccatgttaag 1157
Db 1053 gtaacttggaataatgctatagagaagaacatatatgaataagcctgaagtatcga 1112
Qy 1158 tccctgatgaaagcgttgcttaaaaggtacagaagaaggaattcttgatttaa 1212
Db 1113 aactcaccacaaaagcctctattatagaagatatacgaagagagatagaacgtttaa 1167

```

RESULT 2

```

US-09-302-812-38/c
; Sequence 38, Application US/09302812B
; Patent No. 6333148
; GENERAL INFORMATION:
; APPLICANT: JACOBSON, Myron K.
; APPLICANT: JACOBSON, Elaine L.
; APPLICANT: AM, Jean-Christophe
; APPLICANT: LIN, Winston
; TITLE OF INVENTION: GENES ENCODING SEVERAL POLY(ADP-RIBOSE) GLYCOHYDROLASE (PARC) ENZ
; TITLE OF INVENTION: THE PROTEINS AND FRAGMENTS THEREOF, AND ANTIBODIES IMMUNOREACTIVE
; FILE REFERENCE: N1AD 201
; CURRENT APPLICATION NUMBER: US/09/302,812B
; EARLIER FILING DATE: 1999-04-30
; EARLIER APPLICATION NUMBER: 60/083,768
; NUMBER OF SEQ ID NOS: 38
; SEQ ID NO 38
; LENGTH: 29793
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
; FEATURE:
US-09-302-812-38

```

Query Match 0.9%; Score 90; DB 4; Length 29793;
 Best Local Similarity 50.8%; Pred. No. 5,6e-12;

Matches 309; Conservative 0; Mismatches 275; Indels 24; Gaps 3;

```

Qy 716 actcaaatgaatgatcatgtagtgcgttcaatacaataatttagaattggttgaagc 775
Db 17791 ACCAATGTAATGATGAGCATTCGTCGTCGCCACGCTATCTTATCATTAAGTGAATG- 17731
Qy 776 cgaagacagaatgatcccaacaattcagaagactgtgatgagcgtgcatagctcac 835
Db 17732 -----TTCCGCAATTGGTGAAGACGAGAAAGTCATATTAACATTTGGTCGA 17687
Qy 836 tgaatttgtagatcttgctgcagtgtaagaagcaagccaactggaagtggtg 895
Db 17686 CTAAATCTTGTAAGATTGAGTTGCTTCAGAACGGCAATCGAAACAGAGCAACTGGCGAA 17627
Qy 896 agactaagaagagctgcacaatacgaacgcagctgttacccttggcagagttatga 955
Db 17626 AGATTTAAGAGCTACGAAATCAATCTTCAATTATCGCTCTTGGAATGTATACGT 17567
Qy 956 aagcttagcagcgcagcagcgtggtgattataaactacagacagcaaacaccacca 1015
Db 17566 GCATTGTGGATGCGAATATC---TGCCCATATTCCTTATGAGATCAAAATGATATCGA 17510
Qy 1016 attctcaaatcattcgttggaagaaatgctaaacggttataaattgcacaattaagca 1075
Db 17509 CTCTCTCAAGATTCCTCTCGGTGGAATTCAAAGACAGTATGTTGCTTGTATGAGACT 17450
Qy 1076 g-----ttctttgatgagactctaagctacactcagttgctgagctagctcacaacat 1129
Db 17449 GCAAGTTACATTTTGGAGACACTTGTTACATTAGCATATGCAATCGAAGCTAAAGAAC 17390
Qy 1130 gtgagaataactcccatgttaataagagtcctgagatgaagcgttgctctaaaggttac 1189
Db 17389 ATCAAGATCAACCGAAATTAATGAAGATCCCTAAGGATGCTTACTGCGAGATTCGAA 17330

```

```

Qy 1190 agaagaagaactcttgatattaaagaacaattagagaatttagatcagtcgtgaaca 1249
Db 17329 GAAGAAATTTGAATGCTCCGTGAGCACTGAAAGCAGAGAAAGACAAAGATCGAGGATGA 17270
Qy 1250 aagctcaagcaatgctgaagaagacatacagctgtcagctgaatcacaacta 1309
Db 17269 GCAACTCAAGATTGTTATGATGCTGTAACGACGAACCACTGAAGATGATATGACCAATT 17210
Qy 1310 caacaaga 1317
Db 17209 CAAAAAGA 17202

```

RESULT 3

```

US-09-511-477-38/c
; Sequence 38, Application US/09511477
; Patent No. 6337202
; GENERAL INFORMATION:
; APPLICANT: JACOBSON, Myron K.
; APPLICANT: JACOBSON, Elaine L.
; APPLICANT: AM, Jean-Christophe
; APPLICANT: LIN, Winston
; TITLE OF INVENTION: GENES ENCODING SEVERAL POLY(ADP-RIBOSE) GLYCOHYDROLASE (PARC)
; TITLE OF INVENTION: THE PROTEINS AND FRAGMENTS THEREOF, AND ANTIBODIES IMMUNOREAC
; FILE REFERENCE: N1AD 201
; CURRENT APPLICATION NUMBER: US/09/511,477
; CURRENT FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 09/302,812
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 38
; SEQ ID NO 38
; LENGTH: 29793
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
; FEATURE:
US-09-511-477-38

```

Query Match 0.9%; Score 90; DB 4; Length 29793;
 Best Local Similarity 50.8%; Pred. No. 5,6e-12;

Matches 309; Conservative 0; Mismatches 275; Indels 24; Gaps 3;

```

Qy 716 actcaaatgaatgatcatgtagtgcgttcaatacaataatttagaattggttgaagc 775
Db 17791 ACCAATGTAATGATGAGCATTCGTCGTCGCCACGCTATCTTATCATTAAGTGAATG- 17731
Qy 776 cgaagacagaatgatcccaacaattcagaagactgtgatgagcgtgcatagctcac 835
Db 17732 -----TTCCGCAATTGGTGAAGACGAGAAAGTCATATTAACATTTGGTCGA 17687
Qy 836 tgaatttgtagatcttgctgcagtgtaagaagcaagccaactggaagtggtg 895
Db 17686 CTAAATCTTGTAAGATTGAGTTGCTTCAGAACGGCAATCGAAACAGAGCAACTGGCGAA 17627
Qy 896 agactaagaagagctgcacaatacgaacgcagctgttacccttggcagagttatga 955
Db 17626 AGATTTAAGAGCTACGAAATCAATCTTCAATTATCGCTCTTGGAATGTATACGT 17567
Qy 956 aagcttagcagcgcagcagcgtggtgattataaactacagacagcaaacaccacca 1015
Db 17566 GCATTGTGGATGCGAATATC---TGCCCATATTCCTTATGAGATTCAAACTGACTGGA 17510
Qy 1016 attctcaaatcattcgttggaagaaatgctaaacggttataaattgcacaattaagca 1075
Db 17509 CTCTCTCAAGATTCCTCTCGGTGGAATTCAAAGACAGTATGTTGCTTGTATGAGACT 17450
Qy 1076 g-----ttctttgatgagactctaagctacactcagttgctgagctagctcacaacat 1129
Db 17449 GCAAGTTACATTTTGGAGACACTTGTTACATTAGCATATGCAATCGAAGCTAAAGAAC 17390
Qy 1130 gtgagaataactcccatgttataaggtlccctgagatgaagcgttgctctaaaggttac 1189

```


DB 17389 ATCAAAATCAACCGAATAATATATAGAAATCCCAAGATGCTTTACTGCGAGAAATTCCTCA 17330
QY 1190 agaaagaaatccttgatattaagaacaattagagaatttagagtcagtcgtcgaaca 1249
DB 17329 GAAGAATTTGAATGCTCGCTGAGCAACTGAAGCAGAAAGACAAATGAGATGGA 17270
QY 1250 aaagctcaagcaatgctgaagaagacatagacagtcgtcgtgaatcaacaacta 1309
DB 17269 GCAACTCAAGTTTATGATGCTGAAAGCAGCAAACTTGAGATGATTTGAAGCAAT 17210
QY 1310 cacaaga 1317
DB 17209 CAAAAAGA 17202

RESULT 4
US-08-468-036-23
; Sequence 23, Application US/08468036
; Patent No. 5728806
; GENERAL INFORMATION:
; APPLICANT: Demaggio, Anthony J.
; APPLICANT: Hoekstra, Merl F.
; TITLE OF INVENTION: Materials and Methods Relating to Proteins that
; TITLE OF INVENTION: Interact with Casein Kinase I.
; NUMBER OF SEQUENCES: 48
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468, 036
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/184,605
; FILING DATE: 21-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5728806and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31784
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELE: 25-3856
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5093 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-468-036-23

Query Match 0.9%; Score 88; DB 1; Length 5093;
Best Local Similarity 49.8%; Pred. No. 6.9e-12;
Matches 313; Conservative 0; Mismatches 300; Indels 15; Gaps 3;

QY 610 taatagaaacgtgtatgttgcgtgacactgaagaacttgtaatggttcctgaacatgt 669
DB 1294 TAATCATCATCATCATGATCAAGGGGATGCAGAAATCTTATTAATCTGCACACGAAG 1353
QY 670 aatacagtgatcacaagaagtgtaaaaaaacagacatttgagagagactcaaatgatga 729

DB 1354 CTTGAATTTGCTAATGACGAGGTTGCTTTAAAAAGCAAAAGTGCCCGCTACTAATATCAACGA 1413
QY 730 tcaatagtagtgcgttcacatacaatattagaaatgatttgaaagccgagacaagaatga 789
DB 1414 TCTTTATCATCAAGGCTCTCACACCGCTCTTACAAATCA--CAACAAACATAGTTGAGCANA 1470
QY 790 tcccaaaattcagaagaactgtagtgagagctgtagtgaatcccaattgaatttgtaga 849
DB 1471 TAGCAAAAGCCATGAGCAAAACAAAATTTTGTAAATTTGGCAAAATTTGATTTGGTGG 1530
QY 850 tcttcgtgcagtgtaagaagcaagccaactggaagtgagtgagactlaagaagag 909
DB 1531 TTTGGCAGGACGTGAAGAAACATCAACAGATCGGCTGGGGAATTAAGGCTCAAGAAC 1590
QY 910 ctgcaacatcaaccgacgttgttataccttgagacagtgatataagaagcttagcgag 969
DB 1591 TGCCCTTAATTAACAAATTCGCTTAACACTAGCGCGTGTATCAACGCACT-----CGT 1644
QY 970 ccagagctggtgagattataactacagagacagcaactcagcaattctccaaattc 1029
DB 1645 TGATCATTTCTAACCATATACCTTACAGAGATTAAGCTAACAAATTTGCTACAGACTC 1704
QY 1030 atcggaggaagaatcgcataaagcgtlataatctgcacaaat-----tgcgcagttcttc 1083
DB 1705 TTTAGGTGGATGACGAAGAAACATGATTCATTCGCACTATATACCTCGCAAAATATTCAT 1764
QY 1084 tgatgagactcagtagaacattcagtttgccagtaactgccaaacatgtagaatactcc 1143
DB 1765 GGAAGAGACTGCAAGTACGCTAGCAATATGCAACGAGCCAAATTAATTAAGATCTCC 1824
QY 1144 ccagttaatgagctcgtcgtgagtagaagctggtcgaagaagtagagaagaatctc 1203
DB 1825 ACAAGTAATCACTCTTTATCGAAGATGATATCTTCAGAGACTACATTCAGAGATTTGA 1884
QY 1204 ggaattaaagaacaattagagaattta 1231
DB 1885 AAATTTAAGAAATGATTTGAAAATTTCA 1912

RESULT 5
US-08-376-843-23
; Sequence 23, Application US/08376843
; Patent No. 5846764
; GENERAL INFORMATION:
; APPLICANT: Demaggio, Anthony J.
; APPLICANT: Hoekstra, Merl F.
; TITLE OF INVENTION: Materials and Methods Relating to Proteins
; TITLE OF INVENTION: that Interact with Casein Kinase I
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/376,843
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/184,605
; FILING DATE: 21-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5846764and, Greta E.
; REGISTRATION NUMBER: 35,302
; REFERENCE/DOCKET NUMBER: 27866/31784
; TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 5093 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-376-843-23

Query Match 0.9%; Score 88; DB 2; Length 5093;
Best Local Similarity 49.8%; Pred. No. 6, 9e-12;
Matches 313; Conservative 0; Mismatches 300; Indels 15; Gaps 3;

QY 610 taatagaagaagtgtatgttctgaactgaagaactgttaattgttctgaacatgt 669
DB 1294 TAATTCATCATCATGTCAGGAGATCGAGAAATCTTATTAAGCTGACACGCAAGG 1353
QY 670 aatacagtgatcaaaaagggtgaaaaaacagacattatgagagactaaatgata 729
DB 1354 CTGAATTTCTATGAGGGTTCGTTAAAAGGAAAGTGCCGCTACTAATAAGCAACA 1413
QY 730 tcaatagatcgttcacatacaatattagaaatgattgttgaagcgagacaagaatga 789
DB 1414 TCTTTATCAAGGCTCTCACCCCTTTTACAATCA--CAACAAACATAGTTGACGACAGA 1470
QY 790 tcccaaatctcagagaaactgtgagctgtcagatgatactgaacttaattgtatga 849
DB 1471 TACCAAGACCATGAGCAAAAACAAAATTTGTTAAATTTGCAAAATTTGTTGCGA 1530
QY 850 tcttctgtgcagtgaaagagcaagccaactgagctgaaagggtgagacttaagaag 909
DB 1531 TTTGGCAGGAGTGAACATCAGCAGATCGGTCGCGAGAAATTAAGGGCTCAGGAAGC 1590
QY 910 ctgcacatcaaacccagagctgtttatccttgagacaggtatataagaactagcgacg 969
DB 1591 TGGCCCTAATTAACAAATCCCTCTTAACACTAGGCGGTGTATCAACGCACT-----CGT 1644
QY 970 ccaagctgtgattataactacagagcaagcaaacactcaccagaaattctccaaatctc 1029
DB 1645 TGATCATCTTACCATATACCTTACAGAGAAATCTAAGTACACAAATTTGCTACAGATC 1704
QY 1030 attggaggaagaattgtaaaacggtatataattgcaaat-----tacgcagattctt 1083
DB 1705 TTTAGTGTATGACGAAACATGTCATTATGCAACTATATCACCTGCGAAATATATCAT 1764
QY 1084 tgaatagactcaagtacactcagttgcccagttactccaaacatgtgagaatactcc 1143
DB 1765 GGAAGAGACTGCAACTGAGTATGATGCAACGAGACCCAAATCAATTAAGATTAATCTCC 1824
QY 1144 ccatgttaatgaggtcctcgatgataagcggtgctaaaaagtcagaaagaagaatcct 1203
DB 1825 ACAAGTAAATAGCTTTATGGAAGATACATGCTCAAGACTACATTCAGAGATTGA 1884
QY 1204 ggaattaaagaacaattagaaatta 1231
DB 1885 AAATTTAAGAAATGATTGAAAAATTCGA 1912

RESULT 6
US-08-232-463-14/C
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZpT-Fls
US-08-232-463-14

Query Match 0.9%; Score 85.2; DB 1; Length 7218;
Best Local Similarity 4.5%; Pred. No. 4.1e-11;
Matches 18; Conservative 245; Mismatches 133; Indels 0; Gaps 0;

QY 5271 taaagagaatattgaattgggttgaatttaaaatgaagcgagcaaaagaccacta 5330
DB 1459 TTAAGAGATAGAGAATTTGTCACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1400
QY 5331 aagaacatgtctgtaaatagataaagaacttgagcaagccagcacagactcaat 5390
DB 1399 RRR 1340
QY 5391 gtgagatagaaggtctatagaagacttaagaagatcaagcgctggaacttaaa 5450
DB 1339 RRR 1280
QY 5451 aagaatctgaacaaagaatlaactlgaaccaaagaatgaaatgtttagtcggaaa 5510
DB 1279 RRR 1220
QY 5511 tggaggaattgaaaaatgcccagagactgtlaattgtcgaagagccagctgcaagc 5570
DB 1219 RRR 1160
QY 5571 acctaaagagatgttgaattgtccatltgaactcgaagatcatgaagaagctcaag 5630
DB 1159 RRR 1100
QY 5631 aagcatcagcagcagcaagaataagttcaggaac 5666
DB 1099 RRR 1064

RESULT 7

RESULT 9
US-09-572-191-5
; Sequence 5, Application US/09572191
; Patent No. 6355466
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; APPLICANT: Wood, Kenneth
; TITLE OF INVENTION: No. 6355466el motor proteins and methods for
; FILE REFERENCE: their use
; CURRENT APPLICATION NUMBER: US/09/572,191
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1230
; TYPE: DNA
; ORGANISM: Human
US-09-572-191-5

Query Match 0.9%; Score 82; DB 4; Length 1230;
Best Local Similarity 57.0%; Pred. No. 9.8e-11;
Matches 195; Conservative 0; Mismatches 135; Indels 12; Gaps 2;
QY 830 tctcaattgaattgtgtagatctgtctgagtgaaagagcaagcacaactgagctgaa 889
DB 775 tccctactcaacctggtgatttagcagatcgtgaagaaataatcccatgcaaga 834
QY 890 ggtgtgagacttaagaaagctgcaacataaccgcagctgtttatccttgacaggt 949
DB 835 gggatgagacttgaaagaaagcaagataacatacgaatcgtgagctgctggccaagtg 894
QY 950 attaagaagcttagcagcgccagcgctggtgattataa-----ctacagaagacagc 1003
DB 895 attacagcacttgctgagctggtaatgaaacagagacagttgtctacagagactcc 954
QY 1004 aaactcaccagaattctccaaaattcattgtggagaaatgctaaacggtataattgac 1063
DB 955 aaacttactcttactacaggtatcccttgaggaatgccaacacagccataattgca 1014
QY 1064 acaattacgcagttc-----tttggatgagcttaagtaacatccatggtccagt 1117
DB 1015 aatgttcacccgtgacccagtggtttggggaacccatacacaacttaacttgcctcaa 1074
QY 1118 actgccaacatgtgagaataactcccatgtaataagagtc 1159
DB 1075 agagccaagctgattaaacaaagcagtagtaaatgaagac 1116

RESULT 10
US-09-572-191-1
; Sequence 1, Application US/09572191
; Patent No. 6355466
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; APPLICANT: Wood, Kenneth
; TITLE OF INVENTION: No. 6355466el motor proteins and methods for
; FILE REFERENCE: their use
; CURRENT APPLICATION NUMBER: US/09/572,191
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 4757
; TYPE: DNA
; ORGANISM: Human
US-09-572-191-1

Query Match 0.9%; Score 82; DB 4; Length 4757;
Best Local Similarity 57.0%; Pred. No. 2e-10;
Matches 195; Conservative 0; Mismatches 135; Indels 12; Gaps 2;
QY 830 tctcaattgaattgtgtagatctgtctgagtgaaagagcaagcacaactgagctgaa 889
DB 775 tccctactcaacctggtgatttagcagatcgtgaagaaataatcccatgcaaga 834
QY 890 ggtgtgagacttaagaaagctgcaacataaccgcagctgtttatccttgacaggt 949
DB 835 gggatgagacttgaaagaaagcaagataacatacgaatcgtgagctgctggccaagtg 894
QY 950 attaagaagcttagcagcgccagcgctgtgattataa-----ctacagaagacagc 1003
DB 895 attacagcacttgctgagctggtaatgaaacagagacagttgtctcacaagactcc 954
QY 1004 aaactcaccagaattctccaaaattcattgtggagaaatgctaaacggtataattgac 1063
DB 955 aaacttactcttactacaggtatcccttgaggaatgccaacacagccataattgca 1014
QY 1064 acaattacgcagttc-----tttggatgagcttaagtaacatccatggtccagt 1117
DB 1015 aatgttcacccgtgacccagtggtttggggaacccatacacaacttaacttgcctcaa 1074
QY 1118 actgccaacatgtgagaataactcccatgtaataagagtc 1159
DB 1075 agagccaagctgattaaacaaagcagtagtaaatgaagac 1116

RESULT 11
US-08-480-552-18
; Sequence 18, Application US/08480552
; Patent No. 5665350
; GENERAL INFORMATION:
; APPLICANT: Gudkov, Andrei
; APPLICANT: Roninson, Igor B.
; TITLE OF INVENTION: Genes And Genetic Elements Associated
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 75 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,552
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/033,086
; FILING DATE: 09 MAR 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Keown, Wayne A.
; REGISTRATION NUMBER: 33,923
; REFERENCE/DOCKET NUMBER: 93,354
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/345-9100
; TELEFAX: 617/345-9111
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2389 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA

HYPOTHETICAL: NO
ANTISENSE: YES
US-08-480-552-18

Query Match 0.8%; Score 79.8; DB 1; Length 2389;
Best Local Similarity 49.9%; Pred. No. 4.9e-10;
Matches 430; Conservative 0; Mismatches 377; Indels 54; Gaps 7;

393 gacagacatcttcacagcagacgtacacatgatggaaca-----ccaattcat 443
2 GACAAACATCATCTGGAAGACCCACAGCATGGAGGTTAACTTCATGATCCAGAAGCA 61
444 tggcacaatacccccaagcacaaggaatttcaaatattcagaagat---accga 500
62 TGGGAATTAATTCAGAAAGATAGTCAAGATATTTTAAATTAATTAATTCATGATGATA 121
501 acagagagttcttcataagagttcttataigagagttacataagaatctggaagacc 560
122 ATTTGGAAATTCATATTAAGGTTTCATATTTGAAATATATTTGGATGAATAGAGGACT 181
561 tactgtgatgtacagaagaagacccttgaaatcgcgaagatttcaatagaagc 620
182 TGTATGATGTT-----TCAAAAGACTAACCTTTCAGTCCATGACAGCAAAAACCGTTTC 235
621 tgaatgtgtacgtacgtacgtacgaagactgttaagtgttcctgtaacatgtaacagtgga 680
236 CCTATGTAAGGGGTGCACAGAACCTTTCGTGTAGTCCAGATGAAGTCATGATACCA 295
681 tcaaaaagggtgaaaaaagacacataatgagagactaaatgaatgaatgtagtagc 740
236 TAGATGAAGGGAATCCAAACAGATGTCGAGTTACAAATATGATGAATGAACATACCTCTCA 355
741 gtccacatacaaatattgaatgagttgtgaagcgaagacagaatgaatccacaatc 800
356 GGAGCCACAGCATATTTCTTATTAATGAATAACA-----AGAGATACACAAACCG 406
801 cagagaactgtgagagctgtcatgtatctcactgaattgtgtagactgtctgga 860
407 AACAGAACTCAGTGA-----AAGCTTATCTGTTGATTTAGTGGCA 451
861 gtgaaagagcagaacacgtgagctgaaagtgtgagacttaagaagctgcaacatca 920
452 GTGAAGAGTTAGTAGACTGGGCTGAAGGTGCTGTGATGAATGAAGCAATCA 511
921 accgagctgttataccttgacaggtatataagaagcttagcgaagccagcgtgtg 980
512 AGAAGTCACTTCTGCACTTGGAAATGTCATTTTCGCTTTGGCAGAGGCG-----AGTA 565
981 gattataaactacagagacagaacacacagaatctccaaaatcattctggagaga 1040
566 CCTATGTTCTTATTCAGATATGTAATAATGACCGAATTTCTTCAGATTCATTAAGTGGCA 625
1041 atgcgaagaaggttaatttcacaaatcagcgaagttcttcttgatagagactgaag- 1099
626 ACTGTAGAACCACTATTGTCTATATCTGCTCTCCATCATCATATGATGCTGAGACAA 685
1100 -----acacttcagttccagttactgcaaaacatgtgagaatattccccaagttaag 1154
686 AGTCAACACTCTCTTTGGTGAAGGGCCAAAGCAATTAAGAACACAGCTGTGTCATG 745
1155 aggtcctgtagatgaagcgtgtgtaaaaggtacagaagaagaatcttgatattcaaga 1214
746 TAGAGTTAATCTGACAGAGTGGAAAGAGTATGAAAAAGAAAAAGAAAAAATAAGA 805
1215 aacaatlagagaatttagagt 1235
806 CTCTACGGAACACTATTCAGT 826

RESULT 12
US-08-486-382-4
; Sequence 4, Application US/08486382

Patent No. 5866327
GENERAL INFORMATION:
APPLICANT: Gudkov, Andrei
APPLICANT: Roninson, Igor B.
TITLE OF INVENTION: Association of Kinesin with Sensitivity
TITLE OF INVENTION: To Chemotherapeutic Drugs
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Allegretti & Wilcoff, Ltd.
STREET: 75 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,382
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/177,571
FILING DATE: 05 JAN 1994
ATTORNEY/AGENT INFORMATION:
NAME: No. 5866327/na, Kevin E.
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 93,354-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/715-1000
TELEFAX: 312/715-1234
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2389 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
US-08-486-382-4

Query Match 0.8%; Score 79.8; DB 2; Length 2389;
Best Local Similarity 49.9%; Pred. No. 4.9e-10;
Matches 430; Conservative 0; Mismatches 377; Indels 54; Gaps 7;

393 gacagacatcttcagcagacgtacacatgatggaaca-----ccaattcat 443
2 GACAAACATCATCTGGAAGACCCACAGCATGGAGGTTAACTTCATGATCCAGAAGCA 61
444 tggcacaatacccccaagcacaaggaatttcaaatattcagaagat---accga 500
62 TGGGAATTAATTCAGAAAGATAGTCAAGATATTTTAAATTAATTAATTCATGATGATA 121
501 acagagagttcttcataagagttcttataigagagttacataagaatctggaagacc 560
122 ATTTGGAAATTCATATTAAGGTTTCATATTTGAAATATATTTGGATGAATAGAGGACT 181
561 tactgtgatgtacagaagaagacccttgaaatcgcgaagatttcaatagaagc 620
182 TGTATGATGTT-----TCAAAAGACTAACCTTTCAGTCCATGAAACAAAAACCGTTTC 235
621 tgaatgtgtacgtacgtacgtacgaagactgttaagtgttcctgtaacatgtaacagtgga 680
236 CCTATGTAAGGGGTGCACAGAACCTTTCGTGTAGTCCAGATGAAGTCATGATACCA 295
681 tcaaaaagggtgaaaaaagacacataatgagagactaaatgaatgaatgtagtagc 740
236 TAGATGAAGGGAATCCAAACAGATGTCGAGTTACAAATATGATGAATGAACATACCTCTCA 355
741 gtccacatacaaatattgaatgagttgtgaagcgaagacagaatgaatccacaatc 800

```

Db 356 GGAGCCACGACATATTTCTTATTAATGTAACAA-----AGACAATACAAACG 406
QY 801 cagaagactgtagagcgtgcatgtagtctcactgaattggtagactgtcgga 860
Db 407 AACAGAACTGAGTGA-----AAGCTTTATCTGGTGTGATTTAGTGCA 451
QY 861 gtgaagaagcaagcaaacctggaagctgtagaagctgtagaagcgtgcaatca 920
Db 452 GTGGAAGGTTAGTAAAGACTGGGGCTGAAGTGCTGCTGATGAAGTAAGAACATCA 511
QY 921 accgagctgtttatcctctggacaggtatataagaagcttagcagcgccagctggtg 980
Db 512 AGAAGTACTTTCTGCGACTTGGAATGTCTCTCTGCTTGCGACAGGCG-----AGTA 565
QY 981 gattataaactacagagcaaacactcaccgaattctccaaattatgggagga 1040
Db 566 CCTATGTTCTTATTCGAGATAGTAAATGACCGAATTTCTTCAGATTTCATTAAGTGCA 625
QY 1041 atgctaaagcgttataatttcacaattacgcagttcttcttgtagactctaagt- 1099
Db 626 ACTGTAGGACCACTATTTGTCTATGCTGCTCTCCATCATCATATCAATGAGTCTGAGACA 685
QY 1100 -----acacttcagtttgcagctgccaacatgtagaatactcccaatgtaagt 1154
Db 686 AGTCAACACTCTCTTTGGTCAAAAGGCCAAACAAATTAAGAACACAGTCTGTCGCAATG 745
QY 1155 aggtcctgtagtgaagcgtgtgctaaagaaglacagaagaatcttggatlaaaga 1214
Db 746 TAGAGTTAACTGACAGACGAGTGAAGAAAGACATGATGAAGAAAGAAAAAATAAGA 805
QY 1215 aacaattagagaattagagt 1235
Db 806 CTCTACGGAACACTATTCAGT 826

```

RESULT 13
US-09-235-546-4
; Sequence 4, Application US/09235546
; Patent No. 6043340

```

; GENERAL INFORMATION:
; APPLICANT: Gudkov, Andrei
; APPLICANT: Roninson, Igor B.
; TITLE OF INVENTION: Association of Kinesin with Sensitivity
; TITLE OF INVENTION: No Chemotherapeutic Drugs
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Wilcoff, Ltd.
; STREET: 75 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/235,546
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,382
; FILING DATE:
; APPLICATION NUMBER: US/08/177,571
; FILING DATE: 05 JAN 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6043340nan, Kevin E.
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 93,354-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/715-1000

```

```

; TELEFAX: 312/715-1234
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2389 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEetical: NO
; US-09-235-546-4

```

```

Query Match 0.88; Score 79.8; DB 3; Length 2389;
Best Local Similarity 49.98; Pred. No. 4, 9e-10;
Matches 430; Conservative 0; Mismatches 377; Indels 54; Gaps 7;

```

```

QY 393 gacagacatcttcaggaagcgtgacacatgtagggaaca-----ccaaattcat 443
Db 2 GACAAACATCATCTGGGAGAGCCACACGATGAGGGTTAACTTCATGATCCAGAAAGCA 61
QY 444 tgggcaataaccccacacatacagaagttttaaatlaticagagat---accga 500
Db 62 TGGGAATTAATTCAGAAAGTAGTGCAAGATATTTTAATTAATTAATTCATCCATGATGANA 121
QY 501 acagagagttctctcctaagagttcttatagtagtttacaatgaaactgtgaaagacc 560
Db 122 ATTGGAAATTTCAATATTAAGCTTTCAATATTTGAATTAATTTGATGAAGATAGGACT 181
QY 561 tactgtgtgtagacagaagaagaagcccttggaaattcgcgaggaatttaalagaaga 620
Db 182 TGTATGATGTT-----TCAAGACTAACCTTTCAGCTCCATGAAGCAAAACCGTGTTC 235
QY 621 tgtatgtgtcgtacctgactgtagaagactgttaagtcttccgaatgtaacagtaga 680
Db 236 CCTATGTAAGGGGTGCACAGAACGTTTCTGTGTAGTCCAGATGAAGATGATGATACCA 295
QY 681 tcaaaaaggtgaaaaaaacagacattatgtagagagctaaatgtagatcatagtagtc 740
Db 296 TAGATGAAGGAAATCCAAACAGATGTCGCACTTTCATTAATTAATTAATTAATTAATTA 355
QY 741 gtccacatcaatatttagaatgtagtltgaaagccgagacagaagaatgatcccaaat 800
Db 356 GGAGCCACGACATATTTCTTATTAATGTAACAA-----ACAGAAATACAAACG 406
QY 801 cagaagacgtgtagagctgtagtcatgtagtctcactgaatttggtagactgtcgga 860
Db 407 AACAGAACTGAGTGA-----AAGCTTTATCTGGTGTGATTTAGTGCA 451
QY 861 gtgaagaagcaagccaactgtagcgtgaagctgtagaacttaagaagcgtgcaatca 920
Db 452 GTGGAAGGTTAGTAAAGACTGGGGCTGAAGTGCTGCTGATGAAGTAAGAACATCA 511
QY 921 accgagctgtttatcctctggacaggtatataagaagcttagcagcgccagctggtg 980
Db 512 AGAAGTACTTTCTGCGACTTGGAATGTCTCTCTGCTTGCGACAGGCG-----AGTA 565
QY 981 gattataaactacagagcaaacactcaccgaattctccaaattatgggagga 1040
Db 566 CCTATGTTCTTATTCGAGATAGTAAATGACCGAATTTCTTCAGATTTCATTAAGTGCA 625
QY 1041 atgctaaagcgttataatttcacaattacgcagttcttcttgtagactctaagt- 1099
Db 626 ACTGTAGGACCACTATTTGTCTATGCTGCTCTCCATCATCATATCAATGAGTCTGAGACA 685
QY 1100 -----acacttcagtttgcagctgccaacatgtagaatactcccaatgtaagt 1154
Db 686 AGTCAACACTCTCTTTGGTCAAAAGGCCAAACAAATTAAGAACACAGTCTGTCGCAATG 745
QY 1155 aggtcctgtagtgaagcgtgtgctaaagaaglacagaagaatcttggatlaaaga 1214
Db 746 TAGAGTTAACTGACAGACGAGTGAAGAAAGACATGATGAAGAAAGAAAAAATAAGA 805
QY 1215 aacaattagagaattagagt 1235

```



```

;
; FILING DATE: 09-SEP-1997
; ATTORNEY/AGENT INFORMATION:
;   NAME: Keown, Wayne A.
;   REGISTRATION NUMBER: 33,923
;   REFERENCE/DOCKET NUMBER: 93,354
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 617/345-9111
;   TELEFAX: 617/345-9110
; INFORMATION FOR SEQ ID NO: 18:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 2389 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;     MOLECULE TYPE: cDNA
;     HYPOTHEICAL: NO
;     ANTI-SENSE: YES
; SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-568-315-18

```

Query Match 0.88; Score 79.8; DB 4; Length 2389;

Best Local Similarity 49.9%; Pred. No. 4.9e-10;

Matches 430; Conservative 0; Mismatches 377; Indels 54; Gaps 7;

```

QY 393. gacagacatcttcagcaagaacgtacacaaatgatggaaca-----ccaatcat 443
DB 2 GACAAACATCATCTGGCAAGACCACACGATGAGGGTAACTTCATGATCCAGAGGCA 61
QY 444 tgggcatatcccccaagccatagagaagtttaaatattcagagaat---accga 500
DB 62 TGGGATATTATCCAGATAGTGCAGATATTTTAAATATATTTACTCCATGATGAAA 121
QY 501 acaagagttcttcttaagatttcttataatgagatttacaatgaaactgtgaagacc 560
DB 122 ATTGGAATTTCAATTAAGGTTTCATATTTGAAATATATTGATTAAGATTAAGGACT 181
QY 561 taactgtgtgagacagaagaagaagcccttggaatcgcgagattttaagaaagc 620
DB 182 TGTAGATGTT-----TCAAGACTAACCTTTCAGTCCATGAAGACAAAACCCTGTC 235
QY 621 tgaatgtgctgaactgaactgaactgtaatggttcctgaacatgtaatacagltga 680
DB 236 CCTATGTAAGGGGTGCACAGACGTTTCGTGTAGTCCAGATGAGTCAATGATACCA 295
QY 681 tcaaaaagggtgaaaaaaacagacattatgagagactaaatgatacatagtagtc 740
DB 296 TAGATGAAGGGAAATCCAAACAGAGATGTCGCAATTAATGAATGAATGAACTAGCTCTA 355
QY 741 gtacacatacaatattagaatgattgtgaagccgagacagaatgatcccaaat 800
DB 356 GGAGCCACACGATATTTCTTATTAATGTAACA-----ACAGAAATACACAAACGG 406
QY 801 cagagaactgtgagagctgtcatgtaatcacttgaatttgtagactctgtgca 860
DB 407 AACAGAACTCAGTGA-----AAGCTTATCTGTTGATTTAGCTGCA 451
QY 861 gtgaagaagcaagccaaactgtgaaggtgtgaagacttaagaaaggtcgaacatca 920
DB 452 GTGAGAGGTTAGTAAGACTGGGGCTGAAGGTGCTGTGATGAAGCTAAGAAACATCA 511
QY 921 accgcagctgttatactcttgacagagttatlaagaagcttagcagcgccagctgltg 980
DB 512 AGAAGTCACTTTCTGCACATTGGAAATGTCATTTCTGTTGGCAGAGGSC-----AGTA 565
QY 981 gattataactacagagaacgaacactcaacagaatttcacaaatcattggtgagaa 1040
DB 566 CCTATGTTCTTATCGATAGTAAGTAAGACGAAATTTCTCAAGATTCATTAGGTGGA 625
QY 1041 atgctaaacggttataatgtaattgcaaatgacagcttctttagagactctaagt- 1099
DB 626 ACTGTAGACCACTATTGATATGCTGCTCTCATCATCATATCATATGAGTCTGAGACAA 685

```

```

QY 1100 -----acacttcagtttgccagtaactgccaacaaatgltgagaataactcccaatgtaatg 1154
DB 686 ACTCAACACTCTCTTTGTTCAAAAGGGCCAAATTAAGAACACAGTCTGTCAATG 745
QY 1155 aggtcctgagatgaagagcgttgcctaaaggttacagaagaaggaatcttgatttaaga 1214
DB 746 TAGAGTTAACTCAGAGACAGTGGAAAAAGAGTATGAAAAAGAAAAAATTAAGA 805
QY 1215 aacaaattgagaatttagagt 1235
DB 806 CTCTACGGAACACTATTTCAGT 826

```

Search completed: October 7, 2002, 09:06:54
Job time: 486 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 7, 2002, 08:58:48 ; Search time 9778.42 Seconds
(without alignments)
13286.573 Million cell updates/sec

Title: US-09-150-867-2

Perfect score: 9626
Sequence: 1 gaattccggagtcgtagtag.....ttaaaaaaaaaacgaattc 9626

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estimu:*
5: em_estlov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_estl2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	600.8	6.2	636	10	BM262378 daq40c12.
2	595.8	6.2	681	9	AL638330 AL638330
3	541	5.6	554	9	AM641354 AM641354
4	531.6	5.5	669	10	BJ069784 BJ069784
5	526.4	5.5	583	10	BM262071 BM262071
6	511.4	5.3	545	10	BM262071 daq40c12.
7	453.4	4.7	609	10	BM262071 daq40c12.
8	433.6	4.5	670	9	AL659719 AL659719
9	424.4	4.4	426	10	BI349745 dae63c09.
10	423.8	4.4	457	9	AM636063 BI426c09.w
11	404.4	4.2	612	10	BJ074200 BJ074200
12	378.4	3.9	521	9	AL644957 AL644957
13	371.6	3.9	408	10	BM262071 daq40c12.
14	369.8	3.8	646	10	BM262071 daq40c12.
15	318	3.3	798	9	AM117417 AM117417
16	313.8	3.3	430	10	BM262071 daq40c12.
17	273	2.8	596	9	AL644957 AL644957

18	258.4	2.7	525	9	AA907657
19	254.2	2.6	533	10	UB3037
20	251.8	2.6	430	9	AA964355
21	241	2.5	639	9	AA623883
22	238.4	2.5	546	9	AA811002
23	235.4	2.4	939	10	BE573123
24	234.8	2.4	613	10	BC081419
25	232.6	2.4	1043	10	BE573608
26	221.6	2.3	547	9	AI505210
27	216	2.2	443	9	AA983741
28	203.4	2.1	806	9	AU131640
29	193	2.0	614	10	BU088279
30	189.2	2.0	429	9	AM449147
31	176.2	1.8	409	9	AA721157
32	165.4	1.7	502	10	BE636870
33	165	1.7	376	9	AA287391
34	164.6	1.7	743	10	BI925210
35	161.6	1.7	552	10	BE721105
36	152.6	1.6	757	10	BI933039
37	149.6	1.6	724	10	BI931144
38	149.2	1.5	648	10	BE305901
39	147.8	1.5	533	10	BU086865
40	147.4	1.5	567	10	BM225822
41	147	1.5	557	10	BM228285
42	146.8	1.5	559	10	BM227569
43	146.8	1.5	653	9	AM735873
44	146.8	1.5	1443	11	BC020890
45	143.2	1.5	657	10	BM229569

ALIGNMENTS

RESULT 1
LOCUS BM262378 636 bp mRNA linear EST 18-DEC-2001
DEFINITION daq40c12.Y3 Blackhear/Soares normalized Xenopus egg library
Xenopus laevis cDNA clone IMAGE:4783871 5' similar to TR:042263
042263 KINESIN-RELATED PROTEIN. ; mRNA sequence.

ACCESSION BM262378
VERSION BM262378.1 GI:17925418
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus.
1 (bases 1 to 636)
Clifton,S., Johnson,S.L., Blumberg,B., Song,J., Hillier,L., Pape,D.,
'Martin,J., Wylie,T., Underwood,K., Theising,B., Bowers,Y., Person
,B., Gibbons,M., Harvey,N., Rilter,E., Jackson,Y., McCann,R.,
Waterston,R. and Wilson,R.
Washu Xenopus EST project, 1999
Unpublished (1999)
Other ESTs: daq40c12.x3

TITLE
JOURNAL
COMMENT Contact: Sandy Clifton, Ph.D.
Washu Xenopus EST project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: estelw@wustl.edu
Library constructed by Bento Soares and M. Patricia Bonaldo
(University of Iowa). DNA Sequencing by: Washington University
Genome Sequencing Center
Clone distribution: Xenopus clones from this library are available
through the I.M.A.G.E. Consortium/MLN at: InfocImage.Llnl.gov
Seq primer: -40RP from Gibco
High quality sequence stop: 493.
Location/Qualifiers
1..636
/organism="Xenopus laevis"
/db_xref="taxon:8355"

FEATURES

source

```
/clone="IMAGE:4783871"
/clone_lib="Blackshear/Soares normalized Xenopus egg
library"
/sex="female"
/tissue_type="unfertilized egg"
/cell_type="unfertilized egg"
/dev_stage="unfertilized egg"
/lab_host="DH10B"
/Note="Vector: pT73-Pac; Site_1: EcoRI; Site_2: NotI;
Polya-selected mRNA was prepared from unfertilized Xenopus
laevis eggs. The library was constructed in the vector
pT73-Pac as described in Bonaldo, M.F., Lennon, G. and
Soares, M.B. 'Normalization and subtraction: two
approaches to facilitate gene discovery', Genome Research
6:791-806, 1996. The first strand synthesis used a
NotI-dT18 primer; double stranded cDNAs were ligated to
EcoRI adaptors, digested with NotI, and directionally
cloned into the NotI and EcoRI-digested pT73-Pac vector.
The library contained approximately 7.2 x 105
recombinants, with average insert sizes of 1-1.5 kb."
```

BASE COUNT 185 a 126 c 125 g 200 t

ORIGIN

Query Match 6.2%; Score 600.8; DB 10; Length 636;
Best Local Similarity 99.0%; Pred. No. 1.8e-104;
Matches 616; Conservative 0; Mismatches 2; Indels 4; Gaps 1;

```
QY 8794 gcaaaaagtcacaggaataatcccttaaggaataattatgtatgaagtcaca 8853
    |||||||
DB 15 GCAGAAAGTACAGAAATCTAATCCCTTAAACCAATATTGATGAGTCAGACAA 74
    |||||||
QY 8854 atcgaatgcttaactgcatctcaagtttttataatttaagcttgatgttttcga 8913
    |||||||
DB 75 ATCGATGCTTACTGCTCCATCAGCTTTTGTATTTAGAGCTGGATGATTTTCA 134
    |||||||
QY 8914 getcaacacagagagaaatgacaaagtcagctggaactggtglatgaagcaa 8973
    |||||||
DB 135 GCTCAACACAGAGAGAGATGACAAAGTGCAGGCTGAGTGTGATGAGAGCAA 194
    |||||||
QY 8974 aaaaagaacgacgacgtgaatgtaaaacctagatccctgtaactctgactctcttc 9033
    |||||||
DB 195 AAAAGAAAGCGACCTGAAATGTAACATCTAGATCCCTGATCATCTGACCTCTGTC 254
    |||||||
QY 9034 ctgcaaaagagacttgtaactctgctctctgtaggaagaacacagaacatgcacatg 9093
    |||||||
DB 255 CTGCAAAAGAGACTGCTACTGCTGCTCTGTAAGAGAAACACTAAGAACGCAATGTC 314
    |||||||
QY 9094 tgcataaaggaatcactcctggaagaagaagttgtctttagtaagtaactggttggc 9153
    |||||||
DB 315 TGCATTAAGAGAGTCTCAGTGAAGCAAAAGTTGTTCTTATGTAATCAGTGTGGGC 374
    |||||||
QY 9154 gagggttaacgtcttaataaagtgcaatacgtctcagctctattatattatgtatgt 9213
    |||||||
DB 375 GAGTGGTTACGCTTTAAATTAAGTGAATACGCTCTACGCTTATTTATATGTTAAGT 434
    |||||||
QY 9214 ctgtgtatttgtaactttaagtccttgactcactatatttggctcatctgtatgttct 9273
    |||||||
DB 435 CTGTCTATTGTTTACACTTTTAAAGTCCCTGACTCATATTGCGTCATGCTGATTTCT 494
    |||||||
QY 9274 ttgtgttggcgaacac-----acacactggtgtaaaaaatgacattggcaatgatttlla 9329
    |||||||
DB 495 TTGTGTTTGGCGACACACATACACACTGCTGTAAGAAATACATTGCAATGATGTTTAA 554
    |||||||
QY 9330 ctgactggtctctctggggccatcatatgcaaacacacattggtgcaatggtttttcac 9389
    |||||||
DB 555 CTGACTGCTCTCTCTGGGGCCATCATGCAAAAGCACCATTAGTGTGCCAATGTTTTCAC 614
    |||||||
QY 9390 tacttatattatattgtctgactt 9411
    |||||||
DB 615 TACTTATTATTATATGCTGACTT 636
```

RESULT 2
AL638330 681 bp mRNA linear EST 12-DEC-2001
LOCUS AL638330 XGC-neurula silurana tropicalis cDNA clone TNeu018h15 5',
DEFINITION mRNA sequence.
ACCESSION AL638330.1 GI:16790309
VERSION AL638330.1
KEYWORDS EST.
SOURCE Western clawed frog.
ORGANISM Silurana tropicalis
Euryptera: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
Amphibia: Batrachia: Anura: Mesobatrachia: Pipridae: Pipridae;
Xenopodinae; Silurana.
REFERENCE 1 (bases 1 to 681)
AUTHORS Huckle/E., Taylor/R., Ashurst,J.L., Zorn,A.M. and Rogers,J.
TITLE Sanger Xenopus tropicalis EST project 2001 (10_2001)
JOURNAL Unpublished (2001)
COMMENT Contact: Huckle E
Sanger Centre
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS SEQUENCE ID: TNeu018h15.sp6
Sequencing primer: Sp6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
Location/Qualifiers
.681
/organism="Silurana tropicalis"
/db_xref="taxon:8364"
/clone="TNeu018h15"
/clone_lib="XGC-neurula"
/dev_stage="neurula"
/lab_host="Escherichia coli DH10B"
/Note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
was oligo dt primed from 5ug of poly A+ RNA from neurula.
EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end."

FEATURES
source

BASE COUNT 224 a 129 c 158 g 169 t 1 others

ORIGIN

Query Match 6.2%; Score 595.8; DB 9; Length 681;
Best Local Similarity 93.5%; Pred. No. 1.6e-103;
Matches 632; Conservative 0; Mismatches 43; Indels 1; Gaps 1;

```
QY 100 gggagcggtgtcggtaccagatttccactaatcgttcacaaatgctcgaaggagatgc 159
    |||||||
DB 7 GGGAGCGGTGTGCGGTACCGATTCTGCACCTATGCGACTC-AAATGCCAGGAAGCGC 65
    |||||||
QY 160 agttaagtggtgtgggggttgcgcgttatacagagagaacagggagataagccaa 219
    |||||||
DB 66 AGTAAAGTGTGTGTGAGGTCGCGCGCTTATACAGAGAGCAAGGAGTAAATCCAC 125
    |||||||
QY 220 ccttgcaatggaaaggcttgaaacaacacacattcccaagtgtgatgggacaagtcttcaa 279
    |||||||
DB 126 CCTGCTATWGGAAGCGCTGGAAGCAACACCAATTTCCAGTGTGATGGAGCAAAATCTTTCA 185
    |||||||
QY 280 tttagcagtggtatttaattcactcaagaatacaagaatcaaatlaccagaagaatgagct 339
    |||||||
DB 186 TTTTGATGCTGTATTTAATTCACGAATCAACAAAGCAAGTATTACCAAGAAATATACAGT 245
    |||||||
QY 340 acctatatacgtatcaacttgcagggagataatgycacataattgtcagtaggaagac 399
    |||||||
DB 246 ACCATATATACAGTCACTTTGCAAGGATATATGACCAATATTTTGCTTATGACAGAC 305
    |||||||
QY 400 attctcaggaacagtcacaaatgalyggaacaacaaatltcatltyggcataatgccca 459
    |||||||
DB 306 ATCTTACGCAAGACCTACACGATGATGGAGACCAACCAATTTTGGGCAATATATACCCA 365
    |||||||
QY 460 agccatacagaaggttttaaatltaacagagataacgaaagaagagtttcttctaag 519
    |||||||
DB 366 AGCCATACAGGAAGTGTTAATAATATCCAGAGATACCAACAGAGAGTTCCTTTAAG 425
    |||||||
```

QY 520 agttcttataatggaattacatgaacacgtgaaagaccctactgtgtgatgacgaag 579
DB 426 AGTTCTTATATGAAATTTACATGAACCGTGAAGACCTCCTGTGTGATGACAGAA 485
QY 580 aagaagacccttgaatcgcgagatttaataagaacgtgtatgttctgacctgac 639
DB 486 AAGAAGACCCCTTGAATTCGCGAGGATATCAATGAAGACGTATGTTGCTGACTTAC 545
QY 640 tgaagaactgttaatggttccctgacaatgtaatacagtgatcaaaaaggtgaaaaa 699
DB 546 AGAAGAACCTTGTAATGTTCCGACCATGTAATTCAGTGATCAAAAAGGTCGAAAAA 605
QY 700 cagacattatgagagactaaatgaatgatcatagtagtgcgtccacacattttag 759
DB 606 CAGGATTTATGAGAGACTAAATGATGATCATAGTACCGTCCCATCAATATNTCG 665
QY 760 aatgattgtgaaagc 775
DB 666 AATGATTGTGAAGC 681

RESULT 3
AM641354 554 bp mRNA linear EST 26-APR-2001
LOCUS cm06e04.w1 Blackshear/Soares normalized Xenopus egg library Xenopus
DEFINITION laevis cDNA clone PBX0105E04 5', mRNA sequence.
ACCESSION AM641354
VERSION AM641354.1 GI:7398613
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Xenopus.
1 (bases 1 to 554)
Blackshear, P.J., Lai, W.S., Thorn, J.M., Kennington, E.A., Staffa, N.G.,
Moore, D.T., Bouffard, G.G., Beckstrom-Sternberg, S.M., Touchman,
J.W., Bonaldo, M.F., and Soares, M.B.
The NIHES Xenopus maternal EST project: interim analysis of the
first 13,879 ESTs from unfertilized eggs
Gene 267 (1), 71-87 (2001)
21211403
Contact: Perry J. Blackshear
Office of Clinical Research and Laboratory of Signal Transduction
National Institute of Environmental Health Sciences
A2-05 NIHES, 101 Alexander Drive, Research Triangle Park, NC 27709,
USA
Tel: 919 541-4899
Fax: 919 541-4571
Email: black009@niehs.nih.gov
Clone is available through Research Genetics, Inc., 2130 Memorial
Parkway, Huntsville, AL 35901
phone 800-533-4363 ext.cdna, fax 256-536-9016 att:cdna, email
cdna@resgen.com
DNA sequencing and analyses performed by National Institutes of
Health Intramural Sequencing Center (NISC).
PCR Primers
FORWARD: TGTAAACGACGCCAGT
BACKWARD: CAGGAACAGCATGACG
Plate: 0105 row: E column: 04
Seq primer: 77 primer.
Location/Qualifiers
1. 554
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="PBX0105E04"
/library="Blackshear/Soares normalized Xenopus egg
library"
/sex="female"
/tissue_type="unfertilized egg"
/cell_type="unfertilized egg"
/dev_stage="unfertilized egg"

/lab_host="DH10B"
/note="Vector: pT73-Pac; Site.1: EcoRI; Site.2: NotI;
PolyA-selected mRNA was prepared from unfertilized Xenopus
laevis eggs. The library was constructed in the vector
pT73-Pac as described in Bonaldo, M.F., Lennon, G. and
Soares, M.B. 'Normalization and subtraction', Genome Research
6:791-806, 1996. The first strand synthesis used a
NotI-dT18 primer; double stranded cDNAs were ligated to
EcoRI adapters, digested with NotI, and directionally
cloned into the NotI and EcoRI-digested pT73-Pac vector.
The library contained approximately 7.2 x 10⁵
recombinants, with average insert sizes of 1-1.5 kb."

BASE COUNT 177 a 107 c 129 g 141 t
ORIGIN

Query Match 5.6%; Score 541; DB 9; Length 554;
Best Local Similarity 100.0%; Pred. No. 4.7e-93;
Matches 541; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6962 gctgtgtatcagagcaacacacttgcagcagctctctagtgaagctcaaaagaa 7021
DB 14 GCTGTGTATTGACAGCACACACTTGTCTCAGCAGCTCTCTAGTGAAGCTCAAAAGCA 73
QY 7022 actggaagcacaaacatctgcatgttaataaagaaatctctgcatcagcagctctcc 7081
DB 74 ACTGAAGCACACAAACATTTGCATGTTAAATTAAGAAATCTCTGTCATCCACGCTTCC 133
QY 7082 agatcccttgcagcttgcacaactgcagcagcttgaagcttaataatcacaactgcagaccctt 7141
DB 134 AGATCCTTTGGCAGCTTGCAAACTGACATGTTAACTTAATACATCACTGACAGACCTT 193
QY 7142 ctgaacaaattaaagttgtataccgaactgctgcagtcacgaagaatcatagctgac 7201
DB 194 CTGAACAAATTTAAGGTGTATACCGAATGCTGCAGTCAAGAAAGATCATAGCTTGATC 253
QY 7202 aaagattatgaagaagagacttgcctgcagcaaaagagacagatgagctgacccaa 7261
DB 254 AAGATTAAGAAAGAGACCTTCTGCTGAGCAAAAGAGCAAGATGAGCTGGACCTCCAA 313
QY 7262 ctgcaggttctgagcagcagcagcaaaatgtcggatctgcacctcgaagaaactcaag 7321
DB 314 CTGCAGTGTGTTGGAGCAGCAGCGCAAAATGCTGGATTCGCAATCTGAGAACTCAAG 373
QY 7322 ttctgtgaatgaattcttgaatgagttacttttaaaaaagaacaataatactcaggt 7381
DB 374 TTCTGTGAATTTGAATTTCTTGAATGAGTACTTTTAAAAAAGCAATTAATTCAGAGT 433
QY 7382 gtccaagatgactttcagaagtgcaagttatccctaataatcaagtgatcaaacactgcaa 7441
DB 434 GTCCAGGATGACTTTTCACAGCTGCAGGTATCTTAATAAACAAGTGAAGCAACACTGCAG 493
QY 7442 gaagagcttgagcacaaagaagcttataatgacagtggttgaggaagatttgagatcgac 7501
DB 494 GAAGAGCTTTGACACACAAAGAGCTTTATGACGTGTTGGAGGAATTTGGAGATCTGCAC 553
QY 7502 g 7502
DB 554 G 554

RESULT 4
B069784 669 bp mRNA linear EST 11-DEC-2001
LOCUS B069784 NIBB Mochl1 normalized Xenopus tailbud library Xenopus
DEFINITION laevis cDNA clone XL054124 5', mRNA sequence.
ACCESSION B069784
VERSION B069784.1 GI:17498144
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE	Amphibia: Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus.
AUTHORS	1 (bases 1 to 669) K. Itoyama, A., Terasaka, C., Mochii, M., Ueno, N., Shin-I, T. and Kohara, Y.
TITLE	Expressed genes in <i>X. laevis</i> embryo
JOURNAL	Unpublished (2001)
COMMENT	Contact: Tadasu Shin-I Center For Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshini@genes.nig.ac.jp.
FEATURES	Location/Qualifiers
source	1..669 /organism="Xenopus laevis" /db_xref="taxon:8355" /clone="XLO54124" /clone_lib="NIBB Mochii normalized Xenopus tailbud library" /tissue_type="whole embryo" /dev_stage="stage 25"
BASE COUNT	253 a 122 c 160 g 134 t
ORIGIN	

Query Match	5.5%;	Score 531.6;	DB 10;	Length 669;
Best Local Similarity	88.4%;	Pred No. 36-91;	Indels 3;	Gaps
Matches 589;	Conservative	0;	Mismatches 74;	
QY 7629	gtcaagaatgatgcatgcatgcaacaaggltgaaanaagltgtcaatttgcagaacaactacta	78888		
Db 4	GGCAGAGATGATGCAGGCGCAAAAAAGAAAGAAAGTTGCAATTTTGCAAAAAAACTGCTG	63		
QY 7889	agcgaaagccggaagctgtgctcaaacgcaatgcagtgtaaactactataaagcaagat	79484		
Db 64	AGCGAAACCCGTGAAGCGCGAGCTTAAAGGCAATGAGAGTAAACTGACTGAAAAAGCAAAAC	123		
QY 7949	aatttcaggctgcgaatgaaagaatagaaaactcagaaaatggttgcagaagtgtca	80088		
Db 124	AGTCTTGAGGCTTCATGAAAGAAACCAAAACCTACAGAAAGGTGCCAAAGCTGCA	183		
QY 8009	gtaccatataagaagaagaattgcacaacttaaaactaagtggtgaagaattgaaatgaa	80686		
Db 184	GCACCGGTATTAAGAAAGAAATGTATTAACCTTCAAACCAAGCTGTGAAAGTTGAATGGA	243		
QY 8069	aaaataaagfaccccaagaagaacagacgaagaattgaccttaagaattgtttgaa	81288		
Db 244	AAAATTAAGATTTCGAAAGCAACAGACCAAGAGATTGCTTTTAAAGTCTTGTTGGAA	303		
QY 8129	gataaagaagaagctctgctgtaagttaaaagagaactttagcgagcagcgcaagacac	81888		
Db 304	GATPAGGAAGAAAGGCTGCTGCTAGATTAAAAAGAAACCTTAGACGAGCAAGGCAAGACAC	363		
QY 8189	gatacaacagttgctgttccaaaagttatcagaagaagcttcaactttccctgtgactgt	82488		
Db 364	GATPACCAACATTTGTTTCCAAAAGATTATCGAGAGCTTTCAGTCTCCCTTGGACTTGC	423		
QY 8249	ggtgtgtgaagcggtatagtgcagaagcaagaatgtgtgtcagtcagtcgaaanaagcc	83088		
Db 424	GGTGTGTGAAGCGGCTAGTGTGCGAGACACGCAATGCTGTCTTCAGTGTGAAAAAGCA	483		
QY 8309	gaccttggaaaggaagctgtcacatttaagaagaataatcatcatltaacagactatg	83688		
Db 484	ACGTTGGAAGGAGCGTGTGCGATTATAGAGAAATATCATCATTTTACAGGCTATG	543		
QY 8369	tcaagttctgaa--gacgaaagaanaacaaagcgcaaaatctgatgtcatlcttcat	84258		
Db 544	TCCAAATTTGTAACACATCTTAAGAGAAACAAAGGCAAAATCGATGCGCACTCTTCTAC	603		
QY 8426	accggatcatcaacaagaagctcgctcaacaaactgaacttaagcatcagcattggccctgt	84858		

Db	604	ACTGATATATACACACAGAGGTTACCTTTGCAAAAAGCAAACTTGCACAGCAAGGCCCTGTT	663
Qy	8486	actcca	8491
Db	664	TCTCCA	669

RESULT	5
BM62071/c	
LOCUS	583 bp mRNA linear EST 18-DEC-2001
DEFINITION	dq40c12.x3 Blackshear/Scares normalized Xenopus egg library
Xenopus laevis cDNA clone IMAGE:4783871 3' , mRNA sequence.	
ACCESSION	BM62071
VERSION	BM62071.1 GI:17925111
KEYWORDS	EST.
SOURCE	African clawed frog.

ORGANISM	TITLE	JOURNAL	COMMENT
Xenopus laevis			
Eukaryota: Metazoa: Chordata: Cranata: Vertebrata: Euteleostomi:			
Amphibia: Batrachia: Anura: Mesobatrachia: Piploidea: Pipidae;			
Xenopodinae; Xenopus.			
1 (bases 1 to 583)			
Clifton, S., Johnson, S. L., Blumberg, B., Song, J., Hiller, L., Pape, D.,			
Marlin, J., Wylie, T., Underwood, K., Theising, B., Bowers, Y., Person			
, B., Gibbons, M., Harvey, N., Ritter, E., Jackson, Y., McCann, R.,			
Waterston, R. and Wilson, R.			
Washu Xenopus EST project, 1999			
Unpublished (1999)			
Contact: Sandy Clifton, Ph.D.			

Email: est@watson.wustl.edu
Library constructed by Bento Soares and M. Fatima Bonaldo
(University of Iowa). DNA Sequencing by: Washington University
Genome Sequencing Center
Clone distribution: Xenopus clones from this library are available
through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
Seq Primer: -400P from GIBCO
High quality sequence stop: 510.

```

FEATURES
SOURCE
location/Qualifiers
1..583
/organism="Xenopus laevis"
/db_xref="taxon:8335"
/clone IMAGE:4783871"
/clone_1b="Blackshear/Soares normalized Xenopus egg
library"
/sex="female"
/tissue_type="unfertilized egg"
/cell_type="unfertilized egg"
/dev_stage="unfertilized egg"
/lab_host="DH10B"
/polya_vector: pr773-Pac; Site_1: EcoRI; Site_2: NotI;
NotI-Q718 primer; double stranded cDNAs were ligated to
EcoRI adapters, digested with NotI, and directionally
cloned into the NotI and EcoRI-digested pr773-Pac vector.
The library contained approximately 7.2 X 105
recombinants, with average insert sizes of 1-1.5 kb."
BASE COUNT
201 a 103 c 99 g 180 t

```

Query Match	5.58;	Score 526.4;	DB 10;	Length 583;
Best Local Similarity	98.9%;	Pred. No. 2.9e-90;		
Matches 552; Conservative	0;	Mismatches 1;	Indels 5;	Gaps 2


```

Db      65 AGAAGAGCTTCTTTGAAAAAGATCAGCCTCCAGCAAGCAACTACAGTCAAGGAA 6
Qy      1841 gaaga 1845
Db      5 AAAAA 1

RESULT 7
Bg815925 609 bp mRNA linear EST 22-MAY-2001
LOCUS da075909.v1 Wellcome CRC PCS107 tropicalis egg Silurana tropicalis
DEFINITION cDNA clone IMAGE:44465361 5' similar to TR:042263 042263
KINSHIP-RELATED PROTEIN. ; mRNA sequence.
Bg815925
ACCESSION Bg815925.1 GI:14186905
VERSION EST.
KEYWORDS western clawed frog.
SOURCE Silurana tropicalis
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Silurana.
REFERENCE 1 (bases 1 to 609)
AUTHORS Clifton,S., Johnson,S.L., Blumberg,B., Song,J., Hillier,L., Pape,D.,
Martin,D., Wylie,T., Underwood,K., Theising,B., Bowers,Y., Person
,B., Gibbons,M., Harvey,N., Rilter,E., Jackson,Y., McCann,R.,
Waterson,R. and Wilson,R.
Washu Xenopus EST project, 1999
Unpublished (1999)
TITLE Xenopus EST project, 1999
JOURNAL
COMMENT Contact: Sandy Clifton, Ph.D.
Washu Xenopus EST project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
Library constructed by A. Zorn and J. Mason (Wellcome/CRC Institute
). DNA sequencing by: Washington University Genome Sequencing
Center
Clone distribution: Xenopus clones from this library are available
through the I.M.A.G.E. Consortium/MLN at: info@image.llnl.gov
High quality sequence stop: 467.
FEATURES
Source
1..609
/organism="Silurana tropicalis"
/db_xref="taxon:8364"
/clone_lib="Wellcome CRC PCS107 tropicalis egg"
/tissue_type="egg"
/lab_host="DH10B (phage-resistant)"
/note="Vector: PCS107; Site.1: NotI; Site.2: EcoRI; cDNAs
were oligo-dT primed and directionally cloned. Average
insert size 1.5 kb, range 0.5-4 kb. Library constructed by
A. Zorn and J. Mason (Wellcome/CRC Institute)."
```

```

BASE COUNT 239 a 113 c 129 g 128 t
ORIGIN
```

```

Query Match 4.7%; Score 453.4; DB 10; Length 609;
Best Local Similarity 85.4%; Pred. No. 2; Ce-76;
Matches 505; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
```

```

Qy 5608 agatgatttaagaagctcaagaagcttgacagcagcagaagaataaagttcaggaact 5667
Db 6 AGATGATCTTAAGTTGGACTCAAGAGAATTCACAGCAGCAAGAAATTAAGTTCAAACAGCT 65
Qy 5668 gacctccagattctgtcctgcaggaagaagatctctcttccttggaataatcagatgcttta 5727
Db 66 GACCTCCAGATTCTGTACTGGAAGAAAGATCTCTCTTCTAGAAAATCATGATCTACAGA 125
Qy 5728 taatgttgaactgtgaagaactctaaagcgaagaagaatgactgaaccagcttaagca 5787
Db 126 TAAATGTAACACCAAGAAAGAAATCTCTCAGAGAGATGATCACTGACCAAGTCTAAGCA 185
```

```

Qy 5788 acacgtgtctcagaagaattgaaactctagctgtctttaaagaagaagattgcatt 5847
Db 186 ACACGTGTCTCTGGAGAGGGAAGAACTTAAGCCTGCTTTTAAAGAAAAAGACTTTGCAATT 245
Qy 5848 ggaacagacagagaagaagacaaactgatlgtctgcaggaataacatagatcacagagaa 5907
Db 246 GGAACAAGCAGAAAAGAAAAAGCTGATCAGCCAGCAAGAAATTAATGATCTCATGAGAA 305
Qy 5908 gatataaatatagaagaacagttacttaacaagcaccacatttaagaagaaccttta 5967
Db 306 AATAGCAACTATGAAAGAACAGCTGCTCAAAATGCAACCAATTTTAAAGAAACCTTTTA 365
Qy 5968 tgaagaagaagctctatccagtgtaaggagaacatgctgttgaaacacagaacaccttag 6027
Db 366 TGAANAAGACAGCCTTATTCATCTTAAGAGCAACTGCGCTCGACACTGAACACCTTAG 425
Qy 6028 ggaacattgaagagcaagaacttgcatgggttaaatggggcagggaagagatgaagc 6087
Db 426 GCAAACTTGAAGAGCAAGAACCTGCGATTGGTGAAATGAGACCAAGACAGATGAGAC 485
Qy 6088 tgcacaataaagtatagctctcttaacagaagaatgctctctctagaagaacagatcaatga 6147
Db 486 CGTAATTAAGGTTATGACTCTTACAGAAATATATCTTTCTTGAAGAACAGATCAGTGA 545
Qy 6148 aaatgtactactctaaagaagtgagggtgaanaagagaaccttacct 6198
Db 546 AATGCTACTACTCTTAACAAGTAATGATGAATAATGAAGACCTCAGCCT 596
```

```

RESULT 8
AL659719 670 bp mRNA linear EST 13-DEC-2001
LOCUS AL659719 XGC-neurula Silurana tropicalis cDNA clone TNeu047p06 5',
DEFINITION mRNA sequence.
AL659719
ACCESSION AL659719.1 GI:17673545
VERSION EST.
KEYWORDS western clawed frog.
SOURCE Silurana tropicalis
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Silurana.
```

```

REFERENCE 1 (bases 1 to 670)
AUTHORS Huckle,E., Taylor,R., Ashurst,J.L., Zorn,A.M. and Rogers,J.
TITLE Sanger Xenopus tropicalis EST project 2001
JOURNAL Unpublished (2001)
COMMENT Contact: Huckle E
Sanger Centre
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trops@sanger.ac.uk
```

```

Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: TNeu047p06.sp6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
```

```

FEATURES
Source
1..670
/organism="Silurana tropicalis"
/db_xref="taxon:8364"
/clone_lib="TNeu047p06"
/clone_lib="XGC-neurula"
/dev_stage="neurula"
```

```

/lab_host="Escherichia coli DH10B"
/note="Vector: PCS107; Site.1: EcoRI; Site.2: NotI; cDNA
was oligo dT primed from 5' end of poly A+ RNA from neurula.
EcoRI-NotI cut cDNA was then ligated into PCS107 with
EcoRI at the 5' end and NotI at the 3' end."
```

```

BASE COUNT 218 a 158 c 135 g 159 t
ORIGIN
```

```

Query Match 4.5%; Score 433.6; DB 9; Length 670;
```


Best Local Similarity 84.6%; Pred. NO. 1.6e-72;
Matches 567; Conservative 0; Mismatches 74; Indels 29; Gaps 6;

QY	8247		gtggcgcgaggaacgcgatctatgcagagaacacgcgaactcttctgctcgagtcggaagaag	8306
Dd	6	GTGGAAGGAGGGAAATGTTAATGTCAGACACTGCAGTAATGCTTCTTGCACTCCG--AAAAG	64	
QY	8307	cgcgccttgaaaaaggagcgtgtcacattataagaagaatatcaattatcatacggacta	8366	
Dd	65	CAGCCTTTTGGAAAAAGGAGCTGTGGCATTTTAAACAATAATATCATATTATATACGGACATA	124	
QY	8367	tgtcaagtctcgaa---gatcgaagaagaacaaggccaatactgtatgtcatctcttc	8423	
Dd	125	TATCCAATCTCGAAGATATATCTAAAGAAAACAAGAAGCAAAATCAGATGTGCACCTCTTCC	184	
QY	8424	aatactgatalcatcacacagaaggtctgcctcacaaaacttgaaccttacagacattgscctg	8483	
Dd	185	ATACAGCATATCATCACACAGAGGGTTTTCCCTCACAAAACGTGAACCTTAACAGACATGGCCCCG	244	
QY	8484	ttactccagaagaaggtctgtaaatgcgaagccctacaccttagatctcacagaagaagtcgaagt	8543	
Dd	245	TTTTCTCCAGGAAGGCTCTAAATGCCAAGCCTCAGCCAGATCTCCAAAGAAGTCAAGCT	304	
QY	8544	ccagcactactaaacygttgttgtcaccacaacaggctccgaatactacagccaattagtaagt	8603	
Dd	305	CCCACAAATTAACGTTAGTAGTGTCCACCAAAAGAGTCTGAATGTACAGCCAGTTAGTAATGT	364	
QY	8604	ctccaggaagaagccggagtgat---aaactatctcttctccagaaggtltgagctgc	8660	
Dd	365	CCCCGGCAAGAACCCGGAGTGCATTAATAAACATGTACTTTCTCAAGCAAGGCTGGACCTGC	424	
QY	8661	acaaaagcgtgcccctgctctccaaaacagatcggagatgccacccaagatcgtcatatccc	8720	
Dd	425	ATTAATAACGTAAGACTGTCTCCAAACAGATGTGAGATATCTCATTCAGATGCTGTATGCC	484	
QY	8721	ctggcagaagccggagctgcatataaac-----taactgaagcaagt	8762	
Dd	485	CAGGCAAGCGTGGAGTGCATTAATAATTCACACAGTTTCCCACTCAAGATCAAGGACGACCTC	544	
QY	8763	tatcgacaatttgcctctccatgcaaacagacaagaatlacaggaaaatcctaattccc	8822	
Dd	545	TATTCGGCAATTTG---TCTCATTCGCAAAACGGCAAAATGTACAGSAAATCTAGATTTCC	601	
QY	8823	ctaaagcgaattatttgaigtgaagtcataaaatcgatgccttactgtccatctcagtltt	8882	
Dd	602	CTAATAAC-CATTCTTTTGACATGAAGTCAAAATCTTTCCTATTATGTCATCTCAGTTT	660	
QY	8883	ttgataatic	8892	
Dd	661	TCGATTAATTC	670	
RESULT	9			
B1349745				
LOCUS		426 bp	mRNA	linear
DEFINITION		dbee63c09.y3 Blackshear/Scores normalized Xenopus egg library		
		Xenopus laevis cDNA clone IMAGE:4676672 5' similar to TR-042263		
		O42263 KINESIN-RELATED PROTEIN ; , mRNA sequence.		
ACCESSION				
VERSION		B1349745.1	GI:15044191	
KEYWORDS				
SOURCE				
ORGANISM		African clawed frog.		
		Xenopus laevis		
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
		Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;		
		Xenopodinae; Xenopus.		
REFERENCE		1 (bases 1 to 426)		
AUTHORS		Clifton,S., Johnson,S.L., Blumberg,B., Song,J., Hillier,L., Pape,D.,		
		Martin,J., Wylie,T., Underwood,K., Theising,B., Bowers,Y., Person,		
		B., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R.,		
		Waterston,R. and Wilson,R.		
TITLE		Washu Xenopus EST Project, 1999		
JOURNAL		Unpublished (1999)		

OY 7659 agaaga 7664
 11111
 Db 421 AGAAAA 426

RESULT 10
 AM636063 457 bp mRNA linear EST 26-APR-2001
 LOCUS bl42c09.w1 Blackshear/Soares normalized Xenopus egg library Xenopus
 DEFINITION laevis cDNA clone PBX0042C09 5', mRNA sequence.
 ACCESSION AM636063
 VERSION AM636063.1 GI:7393144
 KEYWORDS EST.
 SOURCE African clawed frog.
 ORGANISM Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 Xenopodinae; Xenopus.
 1 (bases 1 to 457)
 Blackshear, P.J., Lai, M.S., Thorn, J.M., Kennington, E.A., Staffa, N.G.,
 Jr., Moore, D.T., Bouffard, G.G., Beckstrom-Sternberg, S.M., Touchman
 J.W., Bonaldo, M.F. and Soares, M.B.
 The NIH's Xenopus maternal EST project: interim analysis of the
 first 13,879 ESTs from unfertilized eggs
 Gene 267 (1), 71-87 (2001)
 21211403
 Contact: Perry J. Blackshear
 Office of Clinical Research and Laboratory of Signal Transduction
 National Institute of Environmental Health Sciences
 42-05 NIEHS, 101 Alexander Drive, Research Triangle Park, NC 27709,
 USA
 Tel: 919 541-4899
 Fax: 919 541-4571
 Email: black009niehs.nih.gov
 Clone is available through Research Genetics, Inc., 2130 Memorial
 Parkway, Huntsville, AL 35901
 phone 800-533-4363 ext.cdn, fax 256-536-9016 att:cdn, email
 cdna@resgen.com
 DNA Sequencing and analyses performed by National Institutes of
 Health Intramural Sequencing Center (NISC).
 PCR primers
 FORWARD: TGTAAACGACGGCCAGT
 BACKWARD: CAGGAACAGCTATGACC
 Plate: 0042 row: C column: 09
 Seq primer: T7 primer.
 Location/Qualifiers
 1. 457
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /clone="PBX0042C09"
 /clone_lib="Blackshear/Soares normalized Xenopus egg
 library"
 /sex="female"
 /tissue_type="unfertilized egg"
 /cell_type="unfertilized egg"
 /dev_stage="unfertilized egg"
 /lab_host="DH10B"
 /note="Vector: pT73-Pac; Site 1: EcoRI; Site 2: NotI.
 Polya-selected mRNA was prepared from unfertilized Xenopus
 laevis eggs. The library was constructed in the vector
 pT73-Pac as described in Bonaldo, M.F., Lennon, G. and
 Soares, M.B. 'Normalization and subtraction: two
 approaches to facilitate gene discovery', Genome Research
 6:791-806, 1996. The first strand synthesis used a
 NotI-dn18 primer; double stranded cDNAs were ligated to
 EcoRI adapters, digested with NotI, and directionally
 cloned into the NotI and EcoRI-digested pT73-Pac vector.
 The library contained approximately 7.2 x 10⁵
 recombinants, with average insert sizes of 1-1.5 kb."

BASE COUNT 129 a 79 c 74 g 175 t
 ORIGIN

Query Match 4.4%; Score 423.8; DB 9; Length 457;
 Best Local Similarity 98.5%; Pred. No. 1.1e-70;
 Matches 450; Conservative 0; Mismatches 2; Indels 5; Gaps 2;

OY 9155 agtggtaacgtcttaataaagtgcaatacgtctcactcattatattatgtaagtc 9214
 1 AGGTGTTACGTCCTTAAATAAAGCAATACGTCACGTCCTATTATATGTTATGTC 60
 Db 9215 tttgatttggtaacgttttaagtccttgaacttatttggcctcactgattttc 9274
 61 tttgatttggtaacgttttaagtccttgaacttatttggcctcactgattttc 120
 OY 9275 tttgatttggtaacgttttaagtccttgaacttatttggcctcactgattttc 9330
 121 tttgatttggtaacgttttaagtccttgaacttatttggcctcactgattttc 180
 Db 9331 tttgatttggtaacgttttaagtccttgaacttatttggcctcactgattttc 9390
 181 tttgatttggtaacgttttaagtccttgaacttatttggcctcactgattttc 240
 OY 9391 acttatttggtaacgttttaagtccttgaacttatttggcctcactgattttc 9450
 241 acttatttggtaacgttttaagtccttgaacttatttggcctcactgattttc 300
 OY 9451 gtttcctcactgtaacgtatggttacttacttacttacttacttacttactt 9510
 301 gtttcctcactgtaacgtatggttacttacttacttacttacttacttactt 359
 Db 9511 atataatgcatctcaatagtgacctaattgatttacttgcactcttgtaacctttt 9570
 360 atataatgcatctcaatagtgacctaattgatttacttgcactcttgtaacctttt 419
 OY 9571 aaattcactgtaataagcagccctgattttaat 9607
 420 aaattcactgtaataagcagccctgattttaat 456
 Db 420 AAATTTCACCTGTATATAAGCAGCCCTGATTTTAAAT 456

RESULT 11
 BJ074200
 LOCUS BJ074200 612 bp mRNA linear EST 11-DEC-2001
 DEFINITION BJ074200 NIBB Mochii normalized Xenopus tailbud library Xenopus
 laevis cDNA clone XL092h02 5', mRNA sequence.
 ACCESSION BJ074200
 VERSION BJ074200.1 GI:17504389
 KEYWORDS EST.
 SOURCE African clawed frog.
 ORGANISM Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 Xenopodinae; Xenopus.
 1 (bases 1 to 612)
 Kitayama, A., Terasaka, C., Mochii, M., Ueno, N., Shin, I.T. and Kohara
 Y.
 Expressed genes in X. laevis embryo
 Unpublished (2001)
 Contact: Tadashi Shin-I
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin@genes.nig.ac.jp.
 Location/Qualifiers
 1. 612
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /clone="XL092h02"
 /clone_lib="NIBB Mochii normalized Xenopus tailbud
 library"
 /tissue_type="whole embryo"
 /dev_stage="stage 25"

BASE COUNT 191 a 144 c 133 g 144 t

KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
REFERENCE Amphibia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Xenopodinae; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae; 1 (bases 1 to 408)
AUTHORS NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Other-ESTs: dd01b05.y1
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Martha Rebert, Steven L. Klein, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: Xenopus clones from this library are available through the I.M.A.G.E. Consortium/LNL at: info@image.llnl.gov
 Seq primer: -400P from GIBCO
 High quality sequence stop: 320.
FEATURES
 source
 Location/Qualifiers
 1..408
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /clone="IMAGE:3405249"
 /clone_lib="NICHD XGC 001"
 /tissue_type="oocytes"
 /lab_host="DH10B (phage-resistant)"
 /note="Vector: PCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2.2 kb. Constructed by Life Technologies."
BASE COUNT 149 a 62 c 70 g 127 t
ORIGIN
 Query Match
 Best Local Similarity 3.9%; Score 371.6; DB 10; Length 408;
 Matches 399; Conservative 0; Mismatches 4; Indels 5; Gaps 2;
 QY 9220 attgttaacatttaagtccttgactcatatttgctcaatctgtatgttcttgt 9279
 Db 408 ATTGTTACTTAAAGTCCCTGACTCATATTTGGCTCATCTGATGTTCTTGTGT 349
 QY 9280 ttgcgacac-----acacactggtgaaataacatttgcagtgattgttact 9335
 Db 348 TTGGGACACACATACACACTGTGAAAAAATGACATTTGACGTGATTTTACTGACT 289
 QY 9336 ggtctctcggggccatcatgcaaacacattagtgccaaatttttcaactacta 9395
 Db 288 GGCTCTCTGGGGCCATCATGCAACCACTTAGTGCCAAATGTTTTCACCTACTTA 229
 QY 9396 ttattatgctgactcttgaaataacacacaaagatagggcacaagtgtttt 9455
 Db 228 TTAATATGCTGACTTGTGAATAAGAAATATACAAAGATAGGCAAAAGTGTGTT 169
 QY 9456 ctcaataggaacatagtggttaacttaattatctccctaataaatctatata 9515
 Db 168 CTCACCT-GGTAACGATATGGGTTAACTTAATTAATTCCTAATAAATATCTATATA 110
 QY 9516 atggacatctattagtggaactaatttctgcaactctgtaaacctttttaaatt 9575
 Db 109 ATGGCACTCTTATAGTGACCTAATGGTATCTGCACCTCTGTATAAACCCTTTTAAATT 50
 QY 9576 tcaatcgtlaaagcagccatgatttaataaaaaaacaggaa 9623
 Db 49 TCACCTTATAAAGACGCCCTGATTTTAATTAATAAAAAAAA 2
RESULT 14
 BG037357 646 bp mRNA linear EST 24-JAN-2001
 LOCUS BG037357

DEFINITION dc49b03.y1 NICHD XGC Emb3 Xenopus laevis cDNA clone IMAGE:3400276
 5' similar to TR:O42263 O42263 KINESIN-RELATED PROTEIN.; mRNA
ACCESSION BG037357
VERSION BG037357.1 GI:12479942
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
REFERENCE Amphibia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Xenopodinae; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae; 1 (bases 1 to 646)
AUTHORS NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Martha Rebert, Steven L. Klein, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: Xenopus clones from this library are available through the I.M.A.G.E. Consortium/LNL at: info@image.llnl.gov
 Trace considered overall poor quality
 Seq primer: -400P from GIBCO
 High quality sequence stop: 1.
FEATURES
 source
 Location/Qualifiers
 1..646
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /clone="IMAGE:3400276"
 /clone_lib="NICHD XGC Emb3"
 /tissue_type="embryo (stages 24-25)"
 /lab_host="DH10B (phage-resistant)"
 /note="Vector: PCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.7 kb. Constructed by Life Technologies. Note: This is a Xenopus gene collection (XGC) library."
BASE COUNT 181 a 147 c 149 g 169 t
ORIGIN
 Query Match
 Best Local Similarity 3.8%; Score 369.8; DB 10; Length 646;
 Matches 473; Conservative 0; Mismatches 172; Indels 0; Gaps 0;
 QY 4277 aatggccttgaagaagatatttggaacatctgaagaagaagtggtgaagtcacatg 4336
 Db 2 ACTGGCCTTTCACAGGACATATTGGCGAATTTAGACACGCTATGATGAGTCAATG 61
 QY 4337 ttggaagatctgaagaagaacaataagctcaagaagaacagcagaagaatctagt 4396
 Db 62 TTGGAGAACCTGTTGGAGAACCTCAATGACCTCAACGATCAACCAAGGAATATCTAAT 121
 QY 4397 aaagaataatcaatcagtttagaagagtgctcaagtggttcacagaagctgtagatgag 4456
 Db 122 GAATATTCTCAATTCATTAATTTCAATACGTGTTCAAGAGTGGCCAAATCTGCACATGAG 181
 QY 4457 atagaagtcctgaagaacagaagcagaagaagaagcctggaataaagaataga 4516
 Db 182 ATGAGAGATCTGATACGCACTGAAGGACACCTAGAAAGCGCTGGAATTAAGATGAA 241
 QY 4517 gattatttgaacttgtaacaactgcaaacacaaattagtttgaggaagaattggaact 4576
 Db 242 GATTATGATGAACCTGTTCAACGCTGACCCCACTGACAGTCCGGAATATTTGAGAGCT 301
 QY 4577 ccattgcaagctgacacatgagaagaacagcatgtagcggtctggaagaatgagata 4636
 Db 302 ACATTGCAAGCTGCCCATGAGATGACCGTGTGATCAACGTTTTCATCAGCTGGCTATA 361
 QY 4637 aaagtcttgagaagaacttgagcgaatcatcatttacttggaagaattgcaagaaga 4696
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 362 AAGGCTGTGTGAGAGAACTTGACAAATAATTGATCTACTGTCAGAGATTGCTGGAGAA 421
QY 4697 aagctggaactgtctaaacttgaacttcaactcctgaagaagatgagagcgtcgttcta 4756
Db 422 TAGCTGGAAGCTGCTCTCTGCAACTGCTGCTATCCCTTTAAAGAGATGAGACGTGCTG 481
QY 4757 ttaaaagtgtaccttcaacagaagctagaagctgtcgtgagaaacttcttctaaa 4816
Db 482 TCGTAGATCTCTCTCAACATTAACCTGCAACGCTAGCTGAGTGTGCTATCTGATTGTAAT 541
QY 4817 gagaatattgacacacacctaaagcatcatcagatactcaagctcagctgagaaaca 4876
Db 542 GAGATTCTAGAGCGAACCCCTTGACCCCTCCCTCAGAGTGTCTCATGCTCATGCTCAGGCTCA 601
QY 4877 cagcaagagctacagttgtgctaaagaattcttgaatcgtcgttct 4921
Db 602 CTGCAAGCGTACAGATTGCTCTACATCTGCTATCTACTGCTGCT 646

RESULT 15

AU117417 AU117417 798 bp mRNA linear EST 19-OCT-2000
LOCUS AU117417 HEMBA1 Homo sapiens cDNA clone HEMBA1001348 5', mRNA
DEFINITION sequence.

ACCESSION AU117417.1 GI:10932383
VERSION AU117417
KEYWORDS EST.

SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 798)

AUTHORS

Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y.,
Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and
Isogai, T.

TITLE

HRI human cDNA project

JOURNAL

Unpublished (2000)

COMMENT

Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3951
Fax: 81-438-52-3952

Email: genomics@hri.co.jp

HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix
Research Institute; cDNA library construction; Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.

FEATURES

Location/Qualifiers

1..798

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="HEMBA1001348"

/clone_id="HEMBA1"

/tissue_type="whole embryo, mainly head"

/dev_stage="embryo, 10 weeks"

/note="Vector: pM18SF13"

BASE COUNT 262 a 137 c 186 g 209 t 4 others

ORIGIN

Query Match 3.3%; Score 318; DB 9; Length 798;

Best Local Similarity 65.0%; Pred. No. 2e-50;

Matches 506; Conservative 0; Mismatches 257; Indels 15; Gaps 2;

Db 86 aagtcacagagacagagagcggtgtctacagattcttcacatacgtgtcctcaaatg 145

QY 2 AAGTCCCGCGCGGAGGCTCTGCGCATTTTCTGGGACCAAGTTCAGCTCATATGAGATG 61

Db 146 tccagagagagatgcagtttaaggtgtgtgagaggttcggtcgtatatacagagaagacaa 205

QY 62 GCGGAGGAAGAGCGGCTGCTGCTGCGAGTGGCGCGCTGACAGACAGAGAGAGAA 121

QY 206 -----ggggatcaagcccaactgtgcaatggaaagcgtggaacacacacattcccaagt 259
Db 122 TCACCTTGGAGAACTGCCCCAAGTTTACTGGAAACTGACAAATTAATGTCATTATCAAGTT 181
QY 260 gatggacaaagctcttcaatctcagtcggtatcttaacttcaacaaagtcacaa 319
Db 182 GATGGAGAAATTCCTTCAATTTTGTCTGCTTTCATGATGATGAACTACCAAAAT 241
QY 320 attaccagaagatagcagtaacctatcaatcagtcagcttcgaaagatataatggaca 379
Db 242 GTGTATGAAAGATATGACACACCAATCATGATTCCTCATCAAGGCTATACATGCTACT 301
QY 380 atattgtacataggaagacatcttcaagcagaagctacacatgataaggaacacaaat 439
Db 302 ATATTGCTATGAGACAGACTCTTAGGAAACATATACATGATGATGCTTACAGAAAT 361
QY 440 tcaatggacataatcccccaagccatcacaggaagtttcaaaatatacaggaataccg 499
Db 362 CATTTGGAGTTATACCCAGGCGCAATTCATGACATTTTCCAAAATTTAAGAAATTTCTCT 421
QY 500 aacagagagttcttcttaagagttctctatagagatttacaatgaaactgltgaaagac 559
Db 422 GATAGGAATTTCTCTTACGTGATCTTACATGAGAAATATACAAATGAAACCATTTACAGAT 481
QY 560 ctactgtgtgacagaaagaaagagcccttggaaattcgcgagagattttaaagaaac 619
Db 482 TTACTCTGTGGCACTCAAAAATGAACCTTTAATTTTCGAGAAATGTCAATAGCAAT 541
QY 620 gttatgtgtcgaacctgtgaagactgtcaatgttccctgaaacatgtaaacaagttg 679
Db 542 GTGTATGTTGCTGATCTCAGACAGAAAGTTGATATACATCAGAAATGCTTTGAATGA 601
QY 680 atcaaaaaggtgtaaaaaaacagacattatgagagactcaaaatgaaatgacatagtagt 739
Db 602 ATTACAAAGGAGAAAGAGAGGCAATTTAGGAAACAAATGAAATGAAAGAAAGCACT 661
QY 740 cgttcaatacaatatttgaatgttgttgaagcgcgagagaatatacccaaat 799
Db 662 CAGTTCATACCATCTTTAGATGATTTTGAAGAGAGAGAGAGAGAGGCTTAACCTTCTAAT 721
QY 800 tcagaaactgtgagagctgtcagatctcacttgaacttctgtaagatctgctg 857
Db 722 -----TGTGAAGAGATCTGTGAAGNAATCCCATTTGAATTTGGTGTATCTTGCAG 770

Search completed: October 7, 2002, 16:46:41
Job time: 28073 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 7, 2002, 08:58:48 ; Search time 1309.74 Seconds
(without alignments)
12618.549 Million cell updates/sec

Title: US-09-150-867-2

Sequence: 1 gaattcggagtcggatag.....ttaaaaaaaaaaaggaaattc 9626

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_032802:*

- 1: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 11: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 12: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 13: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 14: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 15: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 16: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 17: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 18: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 19: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 20: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 21: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 22: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 23: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
- 24: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	9626	100.0	9626	20	AA26819	Nucleotide sequence
2	620.4	6.4	8083	23	AA570692	DNA encoding novel
3	620.4	6.4	8503	23	AA158253	Human polynucleoti
4	547.6	5.7	8527	22	AA160039	Human polynucleoti
5	206.6	2.1	581	24	AA561800	Lung small cell ca
6	179.4	1.9	10050	23	AB11011	Drosophila melanog
7	141.8	1.5	3440	22	AA52013	Human polynucleoti
8	141.8	1.5	3932	22	AA52997	Human polynucleoti
9	138.2	1.4	501	24	AA561626	Lung small cell ca

C	10	131.2	1.4	12971	23	AB11010	Drosophila melanog
	11	126	1.3	4550	23	AA584836	DNA encoding novel
	12	124.4	1.3	2275	22	AA51140	Human diagnostic a
	13	123.2	1.3	1701	21	AA522377	Human secreted pro
	14	117.6	1.2	3366	21	AA429072	Arabidopsis thalia
	15	107.6	1.1	1950	22	AAH15830	Human cDNA sequenc
	16	106	1.1	814	22	AAH05392	Human cDNA clone (
	17	104.8	1.1	2034	23	AB116733	Drosophila melanog
	18	104.8	1.1	4034	23	AB116732	Drosophila melanog
	19	104	1.1	2352	20	AA487656	Thermomyces lanugl
	20	98.8	1.0	2592	21	AA42909	Arabidopsis thalia
	21	96.4	1.0	7134	23	AB112291	Drosophila melanog
	22	95.8	1.0	1542	22	AA585782	CDNA encoding KSP-
	23	95.8	1.0	1728	22	AA585783	CDNA encoding KSP-
	24	95.8	1.0	3741	22	AAH78013	Nucleotide sequenc
	25	95.8	1.0	3741	22	AA585780	CDNA encoding huma
	26	91.8	1.0	7985	23	AB108391	Drosophila melanog
	27	88	0.9	5093	16	AA099876	KIP1 coding sequen
	28	88	0.9	5093	20	AAV71909	S. cerevisiae KIP1
	29	87.6	0.9	4557	21	AA463311	Arabidopsis thalia
	30	87.6	0.9	4989	21	AA42969	Arabidopsis thalia
	31	86.2	0.9	1107	22	AA585784	CDNA encoding KSP-
	32	85.6	0.9	501	24	AA561778	Lung small cell ca
	33	84.8	0.9	3525	23	AB109775	Drosophila melanog
	34	84.4	0.9	2248	22	AA503053	Human diagnostic a
	35	82	0.9	1230	24	AA23972	Human HSK115 prot
	36	82	0.9	1448	23	AB118979	Drosophila melanog
	37	82	0.9	4757	24	AA233970	Human kinesin supe
	38	82	0.9	4790	22	AA472653	Human cervical can
	39	81.4	0.8	3399	17	AA05868	Chicken leucocytos
	40	79.8	0.8	2389	16	AA52793	Mouse khcs cDNA.
	41	79.2	0.8	1149	22	AA585781	CDNA encoding KSP.
	42	78	0.8	1128	24	AA23971	Human HSK115 prot
	43	77	0.8	3720	23	AB113337	Drosophila melanog
	44	74.4	0.8	463	21	AA47870	Human ORF425
	45	74	0.8	1407	22	AB117294	Human nervous syst

ALIGNMENTS

RESULT 1	AA26819	standard, DNA: 9626 BP.
ID	AA26819	
XX	AA26819:	
AC	AA26819:	
XX	22-JUN-1999	(first entry)
DE	Nucleotide sequence of centromere-associated protein-E (CENP-E).	
XX	CENP-E: centromere-associated protein-E; ATPase activity;	
KW	plus end-directed microtubule motor activity; chromosome congression;	
KW	microtubule binding activity; chromosome movement; mitosis;	
KW	cell proliferation; tumor; metastasis; vascular malfunction;	
KW	inflammatory disease; anglogenesis; hypertension;	
KW	restenosis; fungal infection; selective herbicide; fungicide;	
KW	insecticide; plant growth regulator; activator; cancer cell marker; ss.	
XX		
OS	Xenopus sp.	
XX		
PN	WO9913061-A1.	
XX		
PD	18-MAR-1999.	
XX		
PF	10-SEP-1998; 98WO-US19231.	
XX		
PR	11-SEP-1997; 97US-0058645.	
XX		
PA	(REGC) UNIV CALIFORNIA.	
XX		
PI	Cleveland DM, Goldstein LSB, Sakowicz R, Wood KW;	
XX		

DR WPI; 1999-229233/19.
DR P-PSDB; AAY01632.

Centromere-associated protein-E and related nucleic acid

PS Claim 13; Page 67-73; 77pp; English.

XX The present sequence encodes CENP-E (centromere-associated protein-E)
CC of Xenopus. The protein has at least one of plus end-directed microtubule
CC motor activity, ATPase (adenosine triphosphatase) activity and
CC microtubule binding activity. CENP-E is the motor that powers chromosome
CC movement toward microtubule plus ends and is essential for congression
CC of chromosomes during mitosis. Modulators of CENP-E activity are lead
CC therapeutic, bioagricultural and diagnostic agents, e.g. for treatment
CC of unwanted cell proliferation (typical of many examples are tumors and
CC metastases; vascular malfunction; inflammatory and immune diseases;
CC angiogenesis; hypertension; restenosis; and fungal infections), also as
CC plant-protection agents (selective herbicides, fungicides and
CC insecticides) and plant growth regulators or activators for improving
CC yields. CENP-E is also a diagnostic marker for dividing cells, including
CC cancer cells.

CC Sequence 9626 BP; 3540 A; 1692 C; 2104 G; 2290 T; 0 other;

SQ

Query Match Best Local Similarity 100.0%; Score 9626; DB 20; Length 9626;
Matches 9626; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gaatccggatcgatgagtgatgctggcgaggaatcaactggtatagaagaa 60
Db 1 gaatccggatcgatgagtgatgctggcgaggaatcaactggtatagaagaa 60
QY 61 ctgaaacgcgcgcaaaaggagactaaagtacagagacaggaacggtgctcgatccgat 120
Db 61 ctgaaacgcgcgcaaaaggagactaaagtacagagacaggaacggtgctcgatccgat 120
QY 121 ttccccaactatcggtctcacaagtctccgagggagatgcatggttaaagtgtgtgaggt 180
Db 121 ttccccaactatcggtctcacaagtctccgagggagatgcatggttaaagtgtgtgaggt 180
QY 181 tcggcgcttatacagagagaaacagagggaatcaagccaactgcaatgagagcttgaa 240
Db 181 tcggcgcttatacagagagaaacagagggaatcaagccaactgcaatgagagcttgaa 240
QY 241 caacacattcccaagtgtgatggacaagcttccaattcgatcggtatattaatc 300
Db 241 caacacattcccaagtgtgatggacaagcttccaattcgatcggtatattaatc 300
QY 301 tcaacgaatcaacaagtcaaaatttacaagaagaatagcagtaacctatacgtatcagctt 360
Db 301 tcaacgaatcaacaagtcaaaatttacaagaagaatagcagtaacctatacgtatcagctt 360
QY 361 gcaggagatataatgagcaaatattgcatacagagacatcttcaggcaagctgacac 420
Db 361 gcaggagatataatgagcaaatattgcatacagagacatcttcaggcaagctgacac 420
QY 421 aatgatgggaacacccaattcatcttggtgcaataccccaagccatacagggaagttttaa 480
Db 421 aatgatgggaacacccaattcatcttggtgcaataccccaagccatacagggaagttttaa 480
QY 481 aattatcagagatcacgaacagagagttctctaaagagttcttatataggaattta 540
Db 481 aattatcagagatcacgaacagagagttctctaaagagttcttatataggaattta 540
QY 541 caatgaanaatgtgaaagactactgtgtgatgaacagaaagaccccttggaattcg 600
Db 541 caatgaanaatgtgaaagactactgtgtgatgaacagaaagaccccttggaattcg 600
QY 601 cgaagatttaataagaacgtgtatgttgcgtgactgactgaaagaactctgtaagttcc 660
Db 601 cgaagatttaataagaacgtgtatgttgcgtgactgactgaaagaactctgtaagttcc 660

QY 661 tgaacatgtaatacagtgatgatacaaaagggtgaaaaaacaagacattatgagagactaa 720
Db 661 tgaacatgtaatacagtgatgatacaaaagggtgaaaaaacaagacattatgagagactaa 720
QY 721 aatgatgatcatagtagtgcgttacaatacaataatttagaatgttctggaacccgaga 780
Db 721 aatgatgatcatagtagtgcgttacaatacaataatttagaatgttctggaacccgaga 780
QY 781 caagaatgatcccaaatccagaacgagtgatgagagctgcatatgatacttccacttga 840
Db 781 caagaatgatcccaaatccagaacgagtgatgagagctgcatatgatacttccacttga 840
QY 841 ttggtatgactctgtctgagcagtgaaagacaagccaactgagctgaaagtgtagact 900
Db 841 ttggtatgactctgtctgagcagtgaaagacaagccaactgagctgaaagtgtagact 900
QY 901 taaggagagctgtgaataatacaaccgacgtgtttatcctctggacaggttattaagaagct 960
Db 901 taaggagagctgtgaataatacaaccgacgtgtttatcctctggacaggttattaagaagct 960
QY 961 taagcagcgccaggtctggtgatttaactacagagacagcaaacccacagaaattct 1020
Db 961 taagcagcgccaggtctggtgatttaactacagagacagcaaacccacagaaattct 1020
QY 1021 ccaaaattcattgggaggaatgataaacygtataatttgcaaatlacgcaagtttc 1080
Db 1021 ccaaaattcattgggaggaatgataaacygtataatttgcaaatlacgcaagtttc 1080
QY 1081 ttgtgatgactctcaagacacattcagtttgcaagtcgctgcaaacgctgagaaatlac 1140
Db 1081 ttgtgatgactctcaagacacattcagtttgcaagtcgctgcaaacgctgagaaatlac 1140
QY 1141 tcccatatgttaatgaggtcccttgatgataagcgtttgctaaaggtacagaaaggaat 1200
Db 1141 tcccatatgttaatgaggtcccttgatgataagcgtttgctaaaggtacagaaaggaat 1200
QY 1201 ctgtgatttaagaagaacaatttagagaatttagagttcactcgtctaaacaaagctcagc 1260
Db 1201 ctgtgatttaagaagaacaatttagagaatttagagttcactcgtctaaacaaagctcagc 1260
QY 1261 aatggtctaagaagaagacatacacagttgctagctgaatcaacaactacacaagaagag 1320
Db 1261 aatggtctaagaagaagacatacacagttgctagctgaatcaacaactacacaagaagag 1320
QY 1321 agaagatagatagtcacacttgacaatatgtgttgcctcatcccaagaatctcaaca 1380
Db 1321 agaagatagatagtcacacttgacaatatgtgttgcctcatcccaagaatctcaaca 1380
QY 1381 ggaaccaagggtcaaacgaaacgaagaagtacgctggcgccaggaanaatccaaatatag 1440
Db 1381 ggaaccaagggtcaaacgaaacgaagaagtacgctggcgccaggaanaatccaaatatag 1440
QY 1441 tttaatgcttctgtgttcttctgactttgatatgctataccagattacactggaatttag 1500
Db 1441 tttaatgcttctgtgttcttctgactttgatatgctataccagattacactggaatttag 1500
QY 1501 caagaagaagcaaaagtctcctcagacatgaccttccagaagaatgagatcgtttgtgac 1560
Db 1501 caagaagaagcaaaagtctcctcagacatgaccttccagaagaatgagatcgtttgtgac 1560
QY 1561 agagtttctgatttgaatgacgacctccatcatalgacagaaatgagatagatgacaga 1620
Db 1561 agagtttctgatttgaatgacgacctccatcatalgacagaaatgagatagatgacaga 1620
QY 1621 atggaatttagccagtaaaatcactacagagagaagaagaatcactcacaatctaatgat 1680
Db 1621 atggaatttagccagtaaaatcactacagagagaagaagaatcactcacaatctaatgat 1680
QY 1681 agacttggacagattcttgacagtggttcaagtttcaatgattcttctaaggaanaaccagct 1740
Db 1681 agacttggacagattcttgacagtggttcaagtttcaatgattcttctaaggaanaaccagct 1740

QY 1741 acaatactccccaagaactctgtgatatgctgaatgcagaaaagctctctttgaaaa 1800
|||||
Db 1741 acataactccccaagaactctgtgatatgctgaatgcagaaaagctctctttgaaaa 1800
QY 1801 agagatcaacaagctccccaagaactacagctcaaaaggaaagaaaagaaacttgt 1860
|||||
Db 1801 agagatcaacaagctccccaagaactacagctcaaaaggaaagaaaagaaacttgt 1860
QY 1861 acaaaacttcgagctcaagaatagcagaactcggaagagcgcttaagtgtaaaactaaa 1920
|||||
Db 1861 acaaaacttcgagctcaagaatagcagaactcggaagagcgcttaagtgtaaaactaaa 1920
QY 1921 tcaagagatggttacaaactcgagagagcattccatgaatgctgaagctccaacagatgt 1980
|||||
Db 1921 tcaagagatggttacaaactcgagagagcattccatgaatgctgaagctccaacagatgt 1980
QY 1981 tgaanaaggaagtgtgtgagaanaaagtgtcaagtccttgagagactctgtgtacaaatgcac 2040
|||||
Db 1981 tgaanaaggaagtgtgtgagaanaaagtgtcaagtccttgagagactctgtgtacaaatgcac 2040
QY 2041 aaacagatgacctacagatagctctgttgatgttaaacgtctgaagctcccatgata 2100
|||||
Db 2041 aaacagatgacctacagatagctctgttgatgttaaacgtctgaagctcccatgata 2100
QY 2101 ggtatagaaacacagaanaaagtctggaacaaaagatcggttgatcttagaagatcttga 2160
|||||
Db 2101 ggtatagaaacacagaanaaagtctggaacaaaagatcggttgatcttagaagatcttga 2160
QY 2161 aaaccttaacagaanaaagtgtgagaatgataaacaanaaactctcgagcaaaatcttga 2220
|||||
Db 2161 aaaccttaacagaanaaagtgtgagaatgataaacaanaaactctcgagcaaaatcttga 2220
QY 2221 gagtattcagctatgtgaaagctataatgagcagaanaagcgaactgtgaggaactgac 2280
|||||
Db 2221 gagtattcagctatgtgaaagctataatgagcagaanaagcgaactgtgaggaactgac 2280
QY 2281 acttatgagagaataatttggacaataatttcttagagaatgaaacttaaaagggaat 2340
|||||
Db 2281 acttatgagagaataatttggacaataatttcttagagaatgaaacttaaaagggaat 2340
QY 2341 tgcagatctcggaacgtctcaactcaagagaaatcagaagaaaccaaagtttgaactctga 2400
|||||
Db 2341 tgcagatctcggaacgtctcaactcaagagaaatcagaagaaaccaaagtttgaactctga 2400
QY 2401 gaaggaactcaaaaagaacacagagcacaaactaatacagatgagatctgaaagttaagaa 2460
|||||
Db 2401 gaaggaactcaaaaagaacacagagcacaaactaatacagatgagatctgaaagttaagaa 2460
QY 2461 attagttgaaaaatgcaagatgtaacatcaaaaacttgaggaagatctgaagaactaaac 2520
|||||
Db 2461 attagttgaaaaatgcaagatgtaacatcaaaaacttgaggaagatctgaagaactaaac 2520
QY 2521 aaaactctgaaagagcagagaaatccaactgtcagaaatlaagaaacgcgcgaataact 2580
|||||
Db 2521 aaaactctgaaagagcagagaaatccaactgtcagaaatlaagaaacgcgcgaataact 2580
QY 2581 gcaagaaaaagtacgaataatttgaactctcggttccatgagtgatgagtagtgaagaactctg 2640
|||||
Db 2581 gcaagaaaaagtacgaataatttgaactctcggttccatgagtgatgagtagtgaagaactctg 2640
QY 2641 tgaagaatctttcaacttgaagaactctttcttgatgctcgaagctgtgactgcagatgc 2700
|||||
Db 2641 tgaagaatctttcaacttgaagaactctttcttgatgctcgaagctgtgactgcagatgc 2700
QY 2701 tcaagaagatgtctcttctcctcagaagtgaaaaatcagagctgaagaagaaatctgagaa 2760
|||||
Db 2701 tcaagaagatgtctcttctcctcagaagtgaaaaatcagagctgaagaagaaatctgagaa 2760
QY 2761 cacatcaaacctgttacaatcaaaaagaagcgtcgctcttgtttgagaagcagctgga 2820
|||||
Db 2761 cacatcaaacctgttacaatcaaaaagaagcgtcgctcttgtttgagaagcagctgga 2820
QY 2821 aactgaaaaatcaaaactacaagaanaatggaagctgtattgacagaagaagatgttgaagaagtcg 2880

Db 2821 aactgaaaaatcaaaactacaagaanaatggaagctgtattgacagaagaagatgttgaagaagtcg 2880
|||||
QY 2881 ttttaatgagaatttaactacttaaatgacctcttcggcaggaagaagtcctccagagatctgct 2940
|||||
Db 2881 ttttaatgagaatttaactacttaaatgacctcttcggcaggaagaagtcctccagagatctgct 2940
QY 2941 ttctcggttgaaattgaagaaaaaggtcttcggatcttcgaagcaggtctgaagaagcatt 3000
|||||
Db 2941 ttctcggttgaaattgaagaaaaaggtcttcggatcttcgaagcaggtctgaagaagcatt 3000
QY 3001 ggaagaaaaaaatgaccttggaagaaatgaagtgagacttgctctaaccagaatacaaatcttgcc 3060
|||||
Db 3001 ggaagaaaaaaatgaccttggaagaaatgaagtgagacttgctctaaccagaatacaaatcttgcc 3060
QY 3061 aaatgaagctgaaatgaccttggaanaaaatcagaatcagaagagctctcgaagagaaatgattatc 3120
|||||
Db 3061 aaatgaagctgaaatgaccttggaanaaaatcagaatcagaagagctctcgaagagaaatgattatc 3120
QY 3121 aaagcagaagaggaacatcttcgactatataagcaaaagagatgtatcctgacggn 3180
|||||
Db 3121 aaagcagaagaggaacatcttcgactatataagcaaaagagatgtatcctgacggn 3180
QY 3181 gcaatctgaagcagatcttaacactgagctgaagtgagagtgacacacacagcagcaaaatgga 3240
|||||
Db 3181 gcaatctgaagcagatcttaacactgagctgaagtgagagtgagacacacacagcagcaaaatgga 3240
QY 3241 gcaagctgaaagcaabactctggaatggaanaaatgcatgatactcttctgaaaagta 3300
|||||
Db 3241 gcaagctgaaagcaabactctggaatggaanaaatgcatgatactcttctgaaaagta 3300
QY 3301 tatcagaanaaanaagtgaaagctgaagaccttctaagaagaaatggaagaaacttaagagcac 3360
|||||
Db 3301 tatcagaanaaanaagtgaaagctgaagaccttctaagaagaaatggaagaaacttaagagcac 3360
QY 3361 tatgagatctgtgaaagtaagaatgctgagcacaanaaactgaagactgaaagaaacttaag 3420
|||||
Db 3361 tatgagatctgtgaaagtaagaatgctgagcacaanaaactgaagactgaaagaaacttaag 3420
QY 3421 ggaataaggagcaactgcttcatgagaaaaaatacttcttcaagcaatgagacataat 3480
|||||
Db 3421 ggaataaggagcaactgcttcatgagaaaaaatacttcttcaagcaatgagacataat 3480
QY 3481 tccgattacaacctcttcagagctcgcttccctccaanaattgctgaaaggaaactctca 3540
|||||
Db 3481 tccgattacaacctcttcagagctcgcttccctccaanaattgctgaaaggaaactctca 3540
QY 3541 agaccccatagaataatcaatgactacacaaatttaagacctgtcctcagaagaagagaa 3600
|||||
Db 3541 agaccccatagaataatcaatgactacacaaatttaagacctgtcctcagaagaagagaa 3600
QY 3601 catlatagtgtgctcgaagctgaagaagaaacag tctcagaagagcaagcttaatcttga 3660
|||||
Db 3601 catlatagtgtgctcgaagctgaagaagaaacag tctcagaagagcaagcttaatcttga 3660
QY 3661 caactcaactcaaaagctcttcagacacaaagtatctgaagaatctgatactccagaagccaaa 3720
|||||
Db 3661 caactcaactcaaaagctcttcagacacaaagtatctgaagaatctgatactccagaagccaaa 3720
QY 3721 gcaagacttggaagaggaaggtttaaactgaccttggaatggaacttaactgaaggagca 3780
|||||
Db 3721 gcaagacttggaagaggaaggtttaaactgaccttggaatggaacttaactgaaggagca 3780
QY 3781 cctaactgactcacaagctgtctataagagaatctgcagcttggaagatctggaagtctaca 3840
|||||
Db 3781 cctaactgactcacaagctgtctataagagaatctgcagcttggaagatctggaagtctaca 3840
QY 3841 aaaaactccaaaactctcaagaagaatgaaataatatactcattgaagaagaaatgagcttca 3900
|||||
Db 3841 aaaaactccaaaactctcaagaagaatgaaataatatactcattgaagaagaaatgagcttca 3900
QY 3901 gaccaaactcgaagacttgaaagcagagcgtgaatacctaanaacagaaccttgagaa 3960
|||||

Db	3901	gcccacacttcgaagcctgnaagcgagagcatgatgatagccttaaacaaagaccttagtgaaa	3960
QY	3961	catlgagcagatcaatttgnaacaaacagatgaaatlaagggctgcgcagagaaagcgtaagaga	4020
Db	3961	catlgagcagatcaatttgnaacaaacagatgaaatlaagggctgcgcagagaaagcgtaagaga	4020
QY	4021	acagaagcaactgggttgtaactttatgacaaacagctttatagattgtctctgtaagaaattc	4080
Db	4021	acagaagcaactgggttgtaactttatgacaaacagctttatagattgtctctgtaagaaattc	4080
QY	4081	atcaccaaaccatlgtatgagctgtgttaaaccaagaaaggtgtcatctgggtgtaagtaattc	4140
Db	4081	atcaccaaaccatlgtatgagctgtgttaaaccaagaaaggtgtcatctgggtgtaagtaattc	4140
QY	4141	gttacaagatgaaatlgctgcgtgtgttgaaagagaatgagcttcoaaacatctgtlaagcatt	4200
Db	4141	gttacaagatgaaatlgctgcgtgtgttgaaagagaatgagcttcoaaacatctgtlaagcatt	4200
QY	4201	agtttcgagaaactggagctactctgcgtctgtatgaatctcgttgagaggaaaaaaccttga	4260
Db	4201	agtttcgagaaactggagctactctgcgtctgtatgaatctcgttgagaggaaaaaaccttga	4260
QY	4261	aatcacaaaaaaactcaaatgagccttgaaaaagagatactggcgcaaatctgaagaagaacga	4320
Db	4261	aatcacaaaaaaactcaaatgagccttgaaaaagagatactggcgcaaatctgaagaagaacga	4320
QY	4321	agttgttaagttccactgtcttgtagagatcccaaaaggaaagcaacaataaactaaagaacaagc	4380
Db	4321	agttgttaagttccactgtcttgtagagatcccaaaaggaaagcaacaataaactaaagaacaagc	4380
QY	4381	aaagggaatattctcagtaagaagaaatacaattcagtttagaagaggtgttaagtgttcaca	4440
Db	4381	aaagggaatattctcagtaagaagaaatacaattcagtttagaagaggtgttaagtgttcaca	4440
QY	4441	gaagctgtgtagagatgagataagatcctctgaagaacagcgctaaaggcgcggaagaaagctc	4500
Db	4441	gaagctgtgtagagatgagataagatcctctgaagaacagcgctaaaggcgcggaagaaagctc	4500
QY	4501	ggaaataaaggatagagatattttgaaactgtgacaaactgtgcacaaacacaatlaattagtta	4560
Db	4501	ggaaataaaggatagagatattttgaaactgtgacaaactgtgcacaaacacaatlaattagtta	4560
QY	4561	gggaataatttggaaacactccattgcaacgttgcacatgtaggaagcacaatctgtatcgcgcttc	4620
Db	4561	gggaataatttggaaacactccattgcaacgttgcacatgtaggaagcacaatctgtatcgcgcttc	4620
QY	4621	tgaagaaatvgsgataaaaggtctcttggagaanaaaccttgagcganaatcgaglattaactgga	4680
Db	4621	tgaagaaatvgsgataaaaggtctcttggagaanaaaccttgagcganaatcgaglattaactgga	4680
QY	4681	aagatttgcagaagaagaaagcttggaaactgttcaacaaacttgaatcccttcagaaagagat	4740
Db	4681	aagatttgcagaagaagaaagcttggaaactgttcaacaaacttgaatcccttcagaaagagat	4740
QY	4741	ggagacgttcggttccattaaagaatgagacttgcacaaagagctgagaaagcttgcggagttga	4800
Db	4741	ggagacgttcggttccattaaagaatgagacttgcacaaagagctgagaaagcttgcggagttga	4800
QY	4801	aaacatttcttaaaagagatatgtgacacaacccctaaagcatatttcaagatactcaagc	4860
Db	4801	aaacatttcttaaaagagatatgtgacacaacccctaaagcatatttcaagatactcaagc	4860
QY	4861	tcagcttgcagaanaacacagcagaagctcagatttggcttaagaaattcttgcatacgtctcttc	4920
Db	4861	tcagcttgcagaanaacacagcagaagctcagatttggcttaagaaattcttgcatacgtctcttc	4920
QY	4921	tgcacaattgtccaaatcaactcaagaanaagaaactctgcagatgtgtgcatcctctggga	4980
Db	4921	tgcacaattgtccaaatcaactcaagaanaagaaactctgcagatgtgtgcatcctctggga	4980
QY	4981	agaaagatatgttatttaacttgaagaattgtcatcaaaaaactatgaacaggaanaaat	5040
Db	4981	agaaagatatgttatttaacttgaagaattgtcatcaaaaaactatgaacaggaanaaat	5040

QY	5041	actacatgaaagaatgaaacttgaaagaactcgaagcttgagctaaagctgcagagtggaaaca	5100
Db	5041	actacactgaaagaatgaaacttgaaacaagctcgaagcttgagctaaagctgcagagtggaaaca	5100
QY	5101	ctctgataagagatgatgatgaaatcgaaagccctcatttgatgcttgcgcagatgagaaca	5160
Db	5101	ctctgataagagatgatgatgaaatcgaaagccctcatttgatgcttgcgcagatgagaaca	5160
QY	5161	tgaatctgaaacaacaactctctgctcttaaacagacagaatgcagatgatttactcaagaana	5220
Db	5161	tgaatctgaaacaacaactctctgctcttaaacagacagaatgcagatgatttactcaagaana	5220
QY	5221	gaaagagctgcacacaaccoccatgaaacactttaacagcttgagcttgaccactcaaaagaa	5280
Db	5221	gaaagagctgcacacaaccoccatgaaacactttaacagcttgagcttgaccactcaaaagaa	5280
QY	5281	tatgaaattgggtgttgaaattttaaaaaatgaaacgcagacaaaagaacacactaaagacatg	5340
Db	5281	tatgaaattgggtgttgaaattttaaaaaatgaaacgcagacaaaagaacacactaaagacatg	5340
QY	5341	ctctgataaataagaaataaagaacttgagcgaagccagacacagactccaatgctgaataga	5400
Db	5341	ctctgataaataagaaataaagaacttgagcgaagccagacacagactccaatgctgaataga	5400
QY	5401	agaagcttatgaaagaactttaaagaatgaaagctcgcgcctggaacacttaaaagaattctga	5460
Db	5401	agaagcttatgaaagaactttaaagaatgaaagctcgcgcctggaacactttaaagaattctga	5460
QY	5461	acaaaagaatataatctcttgaaacccaagaatgagaaatggtatgctgcgnaatgaggaatt	5520
Db	5461	acaaaagaatataatctcttgaaacccaagaatgagaaatggtatgctgcgnaatgaggaatt	5520
QY	5521	gaaaaataagccagagagacgtgtaattgctgagagagaccacgtcgaagaacacttaaggga	5580
Db	5521	gaaaaataagccagagagacgtgtaattgctgagagagaccacgtcgaagaacacttaaggga	5580
QY	5581	gaaggttgaaatgtccatttgaaactcaagaatgataagaagaagctccaagaacatttgca	5640
Db	5581	gaaggttgaaatgtccatttgaaactcaagaatgataagaagaagctccaagaacatttgca	5640
QY	5641	gcacgcagaaagaatgaagattcagaagaactgcagacctccagattctgctctccagaaaagt	5700
Db	5641	gcacgcagaaagaatgaagattcagaagaactgcagacctccagattctgctctccagaaaagt	5700
QY	5701	ctctctcttcggaataatcagaatgctctataatgcttgcaactgtgaaagaactctaagcga	5760
Db	5701	ctctctcttcggaataatcagaatgctctataatgcttgcaactgtgaaagaactctaagcga	5760
QY	5761	aagagaatgacctgaaccagctctgaagaacaacctgctctcaagaattgaaactcttagct	5820
Db	5761	aagagaatgacctgaaccagctctgaagaacaacctgctctcaagaattgaaactcttagct	5820
QY	5821	gtctttaaagaanaagaaatttgcatcttggaacaagcagagaaagaaagctgaagctgcgc	5880
Db	5821	gtctttaaagaanaagaaatttgcatcttggaacaagcagagaaagaaagctgaagctgcgc	5880
QY	5881	cagggaaaaaaatagatatcatcacagagagaagatatccaatatagagaagaacttactccaaca	5940
Db	5881	cagggaaaaaaatagatatcatcacagagagaagatatccaatatagagaagaacttactccaaca	5940
QY	5941	agccacccaatttaagaagaaactcttatagaaagagagagctctatccagctgtaagaaaca	6000
Db	5941	agccacccaatttaagaagaaactcttatagaaagagagagctctatccagctgtaagaaaca	6000
QY	6001	actgagctttgaaacacagacaaccttaagggaaacatttgaaagcgaagacttgcatggg	6060
Db	6001	actgagctttgaaacacagacaaccttaagggaaacatttgaaagcgaagacttgcatggg	6060
QY	6061	taaaatgtagcagcagagagatgtagaagctgcgccaataaagataatgactctctacagaagaat	6120
Db	6061	taaaatgtagcagcagagagatgtagaagctgcgccaataaagataatgactctctacagaagaat	6120

QY 6121 gttctcttagaagaacagatcaatgaataatgttactacttaaaagaagttaggttga 6180
|||||
DB 6121 gttctcttagaagaacagatcaatgaataatgttactacttaaaagaagttaggttga 6180
QY 6181 aaaagagaccttactactcaagagaccttcaaaagcagcagctctctcccaatggaga 6240
|||||
DB 6181 aaaagagaccttactactcaagagaccttcaaaagcagcagctctctcccaatggaga 6240
QY 6241 gctcagaagtgctttaagaaccaagaatttgcagttggaagagccgagaagagataag 6300
|||||
DB 6241 gctcagaagtgctttaagaaccaagaatttgcagttggaagagccgagaagagataag 6300
QY 6301 tgaagtgactaatgaataaagaatctcactgttaagaattctctctcttagaagagagat 6360
|||||
DB 6301 tgaagtgactaatgaataaagaatctcactgttaagaattctctctcttagaagagagat 6360
QY 6361 tcttcagaatgtcagacatttgaatgaagctgtgaagcgaagaaggaanaacctgtccatc 6420
|||||
DB 6361 tcttcagaatgtcagacatttgaatgaagctgtgaagcgaagaaggaanaacctgtccatc 6420
QY 6421 gaagcagcaacttggttctcagaatttggagcagctatcactgacatlaaagatlagacca 6480
|||||
DB 6421 gaagcagcaacttggttctcagaatttggagcagctatcactgacatlaaagatlagacca 6480
QY 6481 tgcatttgcgcaatcttaaacgcgaaaagaatgtagctgttaataataatagccagctcgc 6540
|||||
DB 6481 tgcatttgcgcaatcttaaacgcgaaaagaatgtagctgttaataataatagccagctcgc 6540
QY 6541 tgaagaataaagaatctcaacaagaagaatggaatgaattcagaagattcaaaagaacctt 6600
|||||
DB 6541 tgaagaataaagaatctcaacaagaagaatggaatgaattcagaagattcaaaagaacctt 6600
QY 6601 gcaagaacagctctcccatctaaagtgaagaagttagtataataaagacttgaacttcaat 6660
|||||
DB 6601 gcaagaacagctctcccatctaaagtgaagaagttagtataataaagacttgaacttcaat 6660
QY 6661 gctcaagcaacgaaagaagacatcaacaacaacttgcagagaagaagttaaggaataatga 6720
|||||
DB 6661 gctcaagcaacgaaagaagacatcaacaacaacttgcagagaagaagttaaggaataatga 6720
QY 6721 tgaagatctgcaacaacttcatctctaaagagacagcttgagccaatatagatggagct 6780
|||||
DB 6721 tgaagatctgcaacaacttcatctctaaagagacagcttgagccaatatagatggagct 6780
QY 6781 aaggaataaagaagcttcaagaactatgaactctgcgaagaagttagatatacaatggaaaaaga 6840
|||||
DB 6781 aaggaataaagaagcttcaagaactatgaactctgcgaagaagttagatatacaatggaaaaaga 6840
QY 6841 aatctcaagtgctgcgttttaattgagaaagagcctcagcaggaagaatgtagtctgcaga 6900
|||||
DB 6841 aatctcaagtgctgcgttttaattgagaaagagcctcagcaggaagaatgtagtctgcaga 6900
QY 6901 acgtatggtatatacttgagagcagaacaacaagaataacagagctgtagtgaanaaatcttc 6960
|||||
DB 6901 acgtatggtatatacttgagagcagaacaacaagaataacagagctgtagtgaanaaatcttc 6960
QY 6961 cgcgtgtgtatctcaagcaacacacttgcctcagcagctctctctagtgagcttcaaaaga 7020
|||||
DB 6961 cgcgtgtgtatctcaagcaacacacttgcctcagcagctctctctagtgagcttcaaaaga 7020
QY 7021 aacttgaagcaacaacatctgcatgttaataataaagaatctctgtcatccagcgtctc 7080
|||||
DB 7021 aacttgaagcaacaacatctgcatgttaataataaagaatctctgtcatccagcgtctc 7080
QY 7081 cagatctcttgagcagcttgcacaactgagcatgttaagcttaataatcaactgagagccct 7140
|||||
DB 7081 cagatctcttgagcagcttgcacaactgagcatgttaagcttaataatcaactgagagccct 7140
QY 7141 tcttgaacaatttaaggtgttataccgaagctgctcagcttaaaagaagatcatagcttgaat 7200
|||||
DB 7141 tcttgaacaatttaaggtgttataccgaagctgctcagcttaaaagaagatcatagcttgaat 7200
QY 7201 caaagatctatgaaagagacttgctgtgagcaaaagagcatgtagctgtgacttca 7260
|||||
DB 7201 caaagatctatgaaagagacttgctgtgagcaaaagagcatgtagctgtgacttca 7260
QY 7261 actgcagtgcttggagcagcagcagcaaaaatggctggaatctgcacatctgaaggaactcaaa 7320
|||||
DB 7261 actgcagtgcttggagcagcagcagcaaaaatggctggaatctgcacatctgaaggaactcaaa 7320
QY 7321 gttctgtgaattggaattcttgaatgagttactctttaaanaagcaaatataatctcaga 7380
|||||
DB 7321 gttctgtgaattggaattcttgaatgagttactctttaaanaagcaaatataatctcaga 7380
QY 7381 tctcagaagatgacttctcagaagtgagaggtatctccaaatcaagtgagatcaaacactgtca 7440
|||||
DB 7381 tctcagaagatgacttctcagaagtgagaggtatctccaaatcaagtgagatcaaacactgtca 7440
QY 7441 agaaagctctgaagcacaagaagaagcttataagcagtggttggagaagaatttgaagatctgca 7500
|||||
DB 7441 agaaagctctgaagcacaagaagaagcttataagcagtggttggagaagaatttgaagatctgca 7500
QY 7501 cgtcgaatgctaaagaactcagtgaaagcagtgcaacaggaataatagcgcatctgtctctac 7560
|||||
DB 7501 cgtcgaatgctaaagaactcagtgaaagcagtgcaacaggaataatagcgcatctgtctctac 7560
QY 7561 catacagctctttaaacaanaagcgttaagagcagtggttctcagtggaataataacagtgagat 7620
|||||
DB 7561 catacagctctttaaacaanaagcgttaagagcagtggttctcagtggaataataacagtgagat 7620
QY 7621 aaccgtatatactgaacacagctttagaagcgaatattgcaagaagaagaagaacaaacaaga 7680
|||||
DB 7621 aaccgtatatactgaacacagctttagaagcgaatattgcaagaagaagaagaacaaacaaga 7680
QY 7681 acttatgctgagaaatggagcacaagccctagtgcttagtgtaa tggaggaagaataatgc 7740
|||||
DB 7681 acttatgctgagaaatggagcacaagccctagtgcttagtgtaa tggaggaagaataatgc 7740
QY 7741 tagacttttagagatactggaagaacttctcagaatggatcccaagaagaacctccaatcaagagat 7800
|||||
DB 7741 tagacttttagagatactggaagaacttctcagaatggatcccaagaagaacctccaatcaagagat 7800
QY 7801 caaaatgctaaagaatgaaacttgaaacttgcaaaagaatgaaactgcaacaaagtgaaaaa 7860
|||||
DB 7801 caaaatgctaaagaatgaaacttgaaacttgcaaaagaatgaaactgcaacaaagtgaaaaa 7860
QY 7861 agttgcaatttgcagaagacaactactaagcagaagaacgcggaagctgagctaaagcaat 7920
|||||
DB 7861 agttgcaatttgcagaagacaactactaagcagaagaacgcggaagctgagctaaagcaat 7920
QY 7921 gcaagtggaactactcaaaaagaagaataatcttccagcgtgcgaatgaagaagaatgaagaaa 7980
|||||
DB 7921 gcaagtggaactactcaaaaagaagaataatcttccagcgtgcgaatgaagaagaatgaagaaa 7980
QY 7981 cctacagaagaatgcttgcacaagaatgcaagatcacata laaagaagaataatgcaaaccttca 8040
|||||
DB 7981 cctacagaagaatgcttgcacaagaatgcaagatcacata laaagaagaataatgcaaaccttca 8040
QY 8041 aactaagtggtgaagaatggaatggaataaataaagaatctcacaagaacgaacgaaga 8100
|||||
DB 8041 aactaagtggtgaagaatggaatggaataaataaagaatctcacaagaacgaacgaaga 8100
QY 8101 gattgcctacttaagaatcttggttggagaagaatgaaggaagagcttgcgttaggtttaaaga 8160
|||||
DB 8101 gattgcctacttaagaatcttggttggagaagaatgaaggaagagcttgcgttaggtttaaaga 8160
QY 8161 ggaacttaggcagcagcagcagcaacaagaatgaataaagaatgttgcttccaaaagaatataca 8220
|||||
DB 8161 ggaacttaggcagcagcagcagcaacaagaatgaataaagaatgttgcttccaaaagaatataca 8220
QY 8221 gaaagcttcaacttccctgctgactgtgtgtgtgtgaagcgtgtatgtgtagagcagcagc 8280
|||||
DB 8221 gaaagcttcaacttccctgctgactgtgtgtgtgtgtgaagcgtgtatgtgtagagcagcagc 8280
QY 8281 aatgctgtgctgagcttggaagaagccgccttggagaagagctgttcaatataaga 8340
|||||

```

Db 8281 aatgcttgctgctgcaagcgccttggaagggagctgcatattataagaa 8340
QY 8341 gaataatcatcatcttaacacgagctatgtaagttctgaagtcgtaagaagaagaagc 8400
Db 8341 gaataatcatcatcttaacacgagctatgtaagttctgaagtcgtaagaagaagaagc 8400
QY 8401 aaaaatctgctcatctctctcatctgcatcatcacacagagctgctgcctccaaac 8460
Db 8401 aaaaatctgctcatctctctcatctgcatcatcacacagagctgctgcctccaaac 8460
QY 8461 tgaacttaacagacatgagccctgtaactccagaaggtctgaatgcaagcctacact 8520
Db 8461 tgaacttaacagacatgagccctgtaactccagaaggtctgaatgcaagcctacact 8520
QY 8521 aggatctccaaagaagctgagctccagcactaaacgtgttctgacacaaacagctccga 8580
Db 8521 aggatctccaaagaagctgagctccagcactaaacgtgttctgacacaaacagctccga 8580
QY 8581 aatctacagccaattagtaaatgctccagcagaagccggatgcatataactattc 8640
Db 8581 aatctacagccaattagtaaatgctccagcagaagccggatgcatataactattc 8640
QY 8641 tccaagcaaggttgagctgcaaaaagcgtgcccctctccaaacagatcggatgac 8700
Db 8641 tccaagcaaggttgagctgcaaaaagcgtgcccctctccaaacagatcggatgac 8700
QY 8701 caccacagatgcatatccctcctgcaagaccgagctgcatataactaactaagaagc 8760
Db 8701 caccacagatgcatatccctcctgcaagaccgagctgcatataactaactaagaagc 8760
QY 8761 gttatcgaacatttgcctctctcatgcaaacgcaaaaagtaacagtaaatcaattc 8820
Db 8761 gttatcgaacatttgcctctctcatgcaaacgcaaaaagtaacagtaaatcaattc 8820
QY 8821 ccctaagagcaaatatttgaatgtaagtcgaatcgaatgcttactgctcactcgaagt 8880
Db 8821 ccctaagagcaaatatttgaatgtaagtcgaatcgaatgcttactgctcactcgaagt 8880
QY 8881 ttctgataatctcaagctggtgtaatttctcagagctcaacacagcagagagatgacaa 8940
Db 8881 ttctgataatctcaagctggtgtaatttctcagagctcaacacagcagagagatgacaa 8940
QY 8941 aagtcagctgagaaactggtgtaagcaaaaagaagaagcgaacctgtaaaac 9000
Db 8941 aagtcagctgagaaactggtgtaagcaaaaagaagaagcgaacctgtaaaac 9000
QY 9001 atcctagatccctgtaacatctgactcctcctgctcgcgaagaagacttgtaactcgcct 9060
Db 9001 atcctagatccctgtaacatctgactcctcctgctcgcgaagaagacttgtaactcgcct 9060
QY 9061 cttgtagagaagaacaacatgagaactgcatgctcgtcgtcaaaagagtcactgtaagcaa 9120
Db 9061 cttgtagagaagaacaacatgagaactgcatgctcgtcgtcgtcaaaagagtcactgtaagcaa 9120
QY 9121 aagtggtccttagtaagtaactcgtgtggaagtggtgtaagctcttaaaataaagtg 9180
Db 9121 aagtggtccttagtaagtaactcgtgtggaagtggtgtaagctcttaaaataaagtg 9180
QY 9181 caatacgtctcagctcttattatattgtatgctggtgtaattgttacaactttaagtc 9240
Db 9181 caatacgtctcagctcttattatattgtatgctggtgtaattgttacaactttaagtc 9240
QY 9241 ctgactcatatattgctcatctcgtagttcttctgtgttgcaacaacaacgtggtg 9300
Db 9241 ctgactcatatattgctcatctcgtagttcttctgtgttgcaacaacaacgtggtg 9300
QY 9301 aaaaatgacatttgcagtgatatttttactgactggtctcctcgtgggacatcatgcaa 9360
Db 9301 aaaaatgacatttgcagtgatatttttactgactggtctcctcgtgggacatcatgcaa 9360
QY 9361 agcacatagtggtgcaaatgttttcaactactatataatgtcgaacttltgaaata 9420
Db 9361 agcacatagtggtgcaaatgttttcaactactatataatgtcgaacttltgaaata 9420

```

```

QY 9421 gaaaatactacaagaatagggcaaaagtgtttctcactagtgtaaaacagatagtggtt 9480
Db 9421 gaaaatactacaagaatagggcaaaagtgtttctcactagtgtaaaacagatagtggtt 9480
QY 9481 aacttaattatctcctaataaataatctatataatgagcatcattagtgactaat 9540
Db 9481 aacttaattatctcctaataaataatctatataatgagcatcattagtgactaat 9540
QY 9541 gttattcgcgcctcttgcfaaacctttttaaattcactcgttaataagaagccgtat 9600
Db 9541 gttattcgcgcctcttgcfaaacctttttaaattcactcgttaataagaagccgtat 9600
QY 9601 tttaaatataaaaaaaacggaattc 9626
Db 9601 tttaaatataaaaaaaacggaattc 9626

RESULT 2
ID AAS70692 standard; cDNA; 8083 BP.
XX
AC AAS70692;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #6496.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX
OS
XX
PN W0200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001MO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
XX
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
XX
P-PSDB; ABG06505.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
PS Claim 1. SEQ ID NO 6496; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human

```

CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 8083 BP; 3384 A; 1234 C; 1642 G; 1823 T; 0 other;

Query Match 6.4%; Score 620.4; DB 23; Length 8083;
Best Local Similarity 65.5%; Pred. No. 4,9e-131;
Matches 1001; Conservative 0; Mismatches 491; Indels 36; Gaps 5;

QY 69 gccgcgaagagagactaaagtacagagacagggagcggtgttcgtaaccgattccac 128
DB 17 gccgcctgtgaagccctgaagctgcggccggagggctccgcgcaatttctgtggaccg 76
QY 129 taatcgtctcaaaatgtcccgagggagatgcagttaaagtctgtgagaggttcgcgcg 188
DB 77 ttcaagctgatgatgtgctggagagagagcgctgcgtctgcgtgcgagctgcgcgcg 136
QY 189 ttatacagagagacaa-----ggggatcaagccaaacctgcgaatggaagcttgaaaca 242
DB 137 tgaacagcagagaaatcaactctggaagaactgcccaagtttacttgaaacttgacata 196
QY 243 aacacattcccaagttgtatgggacaaagtcttcaattcgaatcgtgtattatctcc 302
DB 197 atgccttattcaagttgatggaagtaaatccctcaattctgactgtcttccatgta 256
QY 303 acgaatcaacagtcacaaattcccaagaataagcaatgacatcatcagacagctttgc 362
DB 257 atgaacactacaaatgtgtatgaagaataagcagcacacatcatcagttcgcgcatc 316
QY 363 agggataataatggtcacataattgtcatcagcagacagatcttcaggaagctgtacaa 422
DB 317 aaggctacataattgtactataattgtcctatggaacagactcttcaggaataacataca 376
QY 423 tgatgggaacaccaaatttcattgtggcataatgcccaagccatagcggaggttttaaa 482
DB 377 tgatgggttcgaagaatcatcttgaggatctataccagggcaattcagatcttccaa 436
QY 483 ttatcagagagataccgaacagagagttctcttaagaattctcttaataagagattaa 542
DB 437 aaattaaagagttctcctgtatggaattctcctcttaagtgatcatcaggaatata 496
QY 543 atgaaactgtgaaagactactgtgtgatgacagaagaagaagcccttgaaattcgcg 602
DB 497 atgaaacccattacagattactctgtgcacatcaaaaaatgaaccttaattatctcg 556
QY 603 aggatttataagaaacggtatgtgtgcagctgactgaagaactgttaattgttcctg 662
DB 557 aagatgtcacaatggaatgtgtatgtgcagctcccaagaagaagcttgatataatcag 616
QY 663 aacatgtcaatcagttgatacaaaaggtgtaaaaaaacagacatcttgagagactaaa 722
DB 617 aaatgcttgaatgtgattacaaagggagaaagagcagcgttatgtgaaacaaaaa 676
QY 723 tgaatgatcatcagtgctgtctacatacatattagaatgattgttgaaagcgagaca 782
DB 677 tgaatcaagaagacgctgcgttcctacatacactcttagagattgttgaaagaagagaga 736
QY 783 gaaatgatcccaaatctcagaagaactgtgatgaggtctcatagttatgtaactttaatt 842
DB 737 aaggtgaaccttcaat-----tgtgaaggaactgttgaaggtatcccatgtgaatt 787
QY 843 tggtagactcttgctgagcagtgaaagagcaagccaaactgagctgaaaggctgtgagacta 902
DB 788 tggtagcttgctgagcagtgaaagagcgtctcaaaacagcgtgaggtgtgtgcgctca 847
QY 903 aggaaggtcgtcaacatcaacgcagctgttataccttvgacaggttataaagaactta 962
DB 848 aggaaggtcgtataataatcgaaagcttattatttggacaagtgatcaagaactta 907
QY 963 ggcagcgccaggtcgtgtgattatataactacagagacgcgaacacaccagatcttcc 1022

DB 908 gtgatgcaagatctgtgtgtttcataataatcogagatagcaagttaaacgaaatcttc 967
QY 1023 aaatctcatctggagagaaatgtcataaacggttataatcttgacaacttaacgcaattct 1082
DB 968 agatctccttggaggaataatccaaagacagatattcttcgacaattatctccagttctt 1027
QY 1083 ctgagtgaactcgaatcagttacacttcagttcgcagttctcccaaacatctgtgaagaatcc 1142
DB 1028 ttgatgaactcttactgtcctcagtttgcagttactgtctaaatataatgaagaatctc 1087
QY 1143 cccatgttaatgaggtcctgagatgaaagcgttgcgttaaaagtctcagaagaagaatcc 1202
DB 1088 ctatgttcaatgagatcaacccgacgaagctccctcgaaaggtatgaaagaataaa 1147
QY 1203 tgaatttaagaagaacatattagaaatttagagatcagtcgtctgaaacaaagctcagaagca 1262
DB 1148 tggattcttaaaaaacatattag-----gaggttctctttagagagcggtgcagcgaa 1201
QY 1263 tggctaaagaagagcatcacagttcgtagctgaaatcaaacacacacacaaagagag 1322
DB 1202 tggaaaaagacccaattggtcccaactcttgaaagaaagaattgtctcgaaggtacaga 1261
QY 1323 aagatagaatagtcacttgacacaaatattgtgtg---ctcatcccaagaatctcaac 1379
DB 1262 atgagaagaattgaaactctacacagagatgctgtgacccctcttccctcagttgcac 1321
QY 1380 aggaacaaaggtgtcaaacgaaacgaagaatlcagtggtgcgcaggaanaaatccaata 1439
DB 1322 aggaataaaggtcaaaacgaagaagttactctgtgtcccttgcaaaatcaaaaaa 1381
QY 1440 gtttaatgctctgtgtgttctgcagtttgatgtgtatcctcagatcttaactcgtgaatttca 1499
DB 1382 tgaagaa-----ctcaactatgcagatcaatlttaataacaaacaaatataa 1429
QY 1500 gcaagaaggaaggtctcgtgacatgctctcaattccagaaatgtgactcgtctgta 1559
DB 1430 caacaaaaacacataagcttctcataaatttatataagagaagaattgtaactcgtctgt 1489
QY 1560 cagagttctcgtatcttgatgacgcct 1587
DB 1490 cagagctgattgttccagtaacacct 1517

RESULT 3
AA158253
ID AA158253 standard; cDNA; 8503 BP.
XX
AC AA158253;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 456.
XX
DE Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX Human; peripheral nervous system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; rheumatic;
XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotoxic;
XX chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX Leukemia; ss.
XX
OS Homo sapiens.
XX
PN WO20015312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000MO-US34263.
XX
PR 21-JAN-2000; 2000US-0489725.
XX
PR 25-APR-2000; 2000US-0552317.
XX
PR 09-JUL-2000; 2000US-0598042.
XX
PR 19-JUL-2000; 2000US-0620312.

RESULT 4
AA160039/C
ID AA160039 standard; CDNA; 8527 BP.
AC AA160039;
XX 22-OCT-2001 (first entry)
DT
XX Human polynucleotide SEQ ID NO 4028.
DE
XX
XX Human; nocotropic; immunosuppressant; cytoskeletal; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
XX Homo sapiens.
OS
XX WO200153312-A1.
PN
XX 26-JUL-2001.
PD
XX 26-DEC-2000; 2000WO-US34263.
XX
XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
PA
XX Tang YF, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI Zhao Q, Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI: 2001-442253/47.
XX P-PSDB: AAM40883.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders
PT as central nervous system injuries -
XX
PS Claim 1; SEQ ID NO 4028; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and
CC the encoded polypeptides (AAM38642-AAM42213) with nocotropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localized neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilization of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokine activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders.
CC Note: The sequence data for this patent did not form part of the printed
CC specification.
XX
SQ Sequence 8527 BP; 1982 A; 1715 C; 1302 G; 3525 T; 3 other;

Query Match 5.7%; Score 547.6; DB 22; Length 8527;
Best Local Similarity 64.2%; Pred. No. 2e-114;
Matches 988; Conservative 0; Mismatches 504; Indels 48; Gaps 9;

QY 69 gccgcacaaaggaactaaagtgacagagagagagcggtgtcgtgtacccgattccccc 128
DB 8511 GCGGCTGTGAGCCCTGAATGCTCCGCCGCCGAGGCTCTGGCATTTTGTGGACAC 8452
QY 129 taatcgctcacaatgtccgaagagatgacgttaagtgtgtgtgaggttcggccgc 188
DB 8451 TTCAGCCTGATGATGATGCGGAGAGAGACCCTGGCCCTCTGCGAGTGGCCGC 8392
QY 189 ctatcagaagagacaa-----gggatcacaagccaactgtcaatggaagctgaaaca 242
DB 8391 TGAACAGACAGAGAAGATCACTTGAGAAACCTGCCCAAGTTTACTGMAAACACCAATA 8232
QY 243 acacacattcccaagtgtatgaggaacaagtcttcaattcgtgtgttcaattc 302
DB 8331 ATGTCATTTATCCAGTGAAGTAATAAATCCCTTCAATTTTGTGCTGCTTCATGTA 8272
QY 303 acgaatcaacaagctcaatttaccag-aatatagcaatcctatcaatcagctgtg 361
DB 8271 ATGAACCTCCCAAAATGTGTATGAAGCCATATGACACACCATCATGATTCGCCATA 8212
QY 362 caggatataatgycacaatattg-catacggacagac-alcitcaggaagacgtaca 419
DB 8211 CAAGCTACAAATGATATATTGCTGATGAGACAGACTGCTTCAGGAAAAACATATA 8152
QY 420 caatgttggaacaacaaatctctgtgcaatlaaccccaagcacaacaggaagtttc- 478
DB 8151 CCATGATGGCTTCAGAAATCATATGAGATTTGGAGTATACCCCGGCAATTTATGATG 8092
QY 479 -----aaattatcagagataccgaagagagttcttcgaagttctcata 530
DB 8091 TTTCACAAAATTTTAAGAAAGTTTCTCTGATAGGGAATTTCTTACGATCTTACA 8032
QY 531 tggagattcaaatgaactgtgaagaccctactgtgtgacagagagaagaagccct 550
DB 8031 TCGAAATATATCAATGAACCATTAACAGATTACTGTGACATCAAAAATGAAACCTT 7972
QY 591 tggaaattcgagagatttaataagaacgtgtatgtctgtccgcgcgaagaactg 650
DB 7971 TAATTTATCGAGAGATGTCMAATAGGAAATGTAATTTGCTGTATTCACAGAAAGTTC 7912
QY 651 taatgtctctgacacatgtaatacagtgagtcataaagaagtgaaaaaagaacatla 710
DB 7911 TATATCATCAGAAATGCGCTTTCAAAATGATTCMAAGGAGAAAGACGACCATTTATG 7852
QY 711 gagagactaaatgaatgacatagtagtgcgtcacaataatltagaatgtgtg 770
DB 7851 GAGAAACAAATATGATCAAGAAAGAGAGAGCGTTCATACCATCTTAGATGATTTGG 7792
QY 771 aaagccgagacagaatgatccacaataatcagagaactgtgacgtgtcgtat 830
DB 7791 AAAGCAAGAGAGAGGCTGTAACCTTTCTAAAT-----TGTGAAGATCTGTTAAGCTAT 7741
QY 831 ctacactgaattgtgtagtctgtcgtgagtgaaagacaaacgtgagtcgaag 890
DB 7740 CCATTTGAATTTGTTGATCTTGTGACGCGTAAAGAGCTGCTCAAAAGGGGCTGCAG 7681
QY 891 gtcgtgaacttaaggaagctgtcaacaacaacgcagactgttatacctctgacagtt 950
DB 7680 GTGTGGGCTCAAGAGAGGCTGTAAATATATGAAAGCTTATTTTGGGCAAGTGA 7621
QY 951 ttaagaacttagcagagccaggtcgtgtgatttataactacagagacagaactca 1010
DB 7620 TCAGAAACTTACTAGTGTGAGCAAGTGTGCTTCAATTAATTTCAAGAACACACTTAA 7561
QY 1011 ccagaattctcaaaatctaatgtgaggaatgtcaaacggtatacaattgtcaactta 1070
DB 7560 CACGAATTTCTCAGAAATCTTGTGGAGAAATCAAGAACACGATATATCTGCACATTA 7501
QY 1071 cgcgaattcttltgtgagactcctaagtaactcagtttgcagtaactgtcacaacag 1130
DB 7500 CTCACATATCTTTTGTATGAAGACTTACTGCTCTCCAGTTTGGCATATACGCTAAATATA 7441

CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAK80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.
 CC
 XX

Sequence 3932 BP; 1312 A; 771 C; 913 G; 936 T; 0 other;

Query Match 1.5%; Score 141.8; DB 22; Length 3932;
 Best Local Similarity 52.8%; Pred. No. 5.5e-22;
 Matches 504; Conservative 0; Mismatches 397; Indels 54; Gaps 7;

QY 263 ggagacaagctctcaattcagctgctgattatcctcagaatcaacaagcaatt 322
 DB 558 ggcgcacagccttatgcatctgacggtggtccagtcacacacccacagcagtg 617
 QY 323 taccaagaataagcagtaacctacatcacatcagctcttgacggatataatgacaata 382
 DB 618 tataatgactgtgcagaagaagtgttaagaatgactcttgaaagatatatgacaata 677
 QY 383 ttgtcacaagcagacacaccttcagcagaagcgtacacaatgattggagaac----- 434
 DB 678 ttgtcatalgacacaacacacctctggagaaacacacacatggaggttaactcatgat 737
 QY 435 -caaatcatctggacatacccaagccatcaggaagtcttaaatattcaggag 493
 DB 738 ccaggaagcagatgggaattctccagaataagtcagaataattcttaattatctcc 797
 QY 494 at--accgaacagagattctcttaagaagttcttataatgagataaactgaact 550
 DB 798 atgagatgaataatttgaaattcatttaaggtttcattatgtaataatatttgagtaag 857
 QY 551 gtagaagaccctactgtgtatgacagaagaagaagcccttgaaattcgagagatttc 610
 DB 858 ataagagaccctgtatagttt-----tcaagaaccaaccttcagttcctgaagacaa 911
 QY 611 aatagaacagtgatgtctgacccctgacgaagaacttgtaagtcttcctgacatgta 670
 DB 912 aaccgaagtcctcattatgaaagggtgcacagacgcttctgtatgctcagatgaagt 971
 QY 671 atacagtgatcaaaaagggtgaaaaaacaagacattatgagagacataaagaattgat 730
 DB 972 atgagatcacatagatgaagaaatccacaagacatgtagcattacaatatagatgaa 1031
 QY 731 catagtagtcgttcaatatacatattatgattgttgaaagccgagacagaatgat 730
 DB 1032 catagctcaggagatcagcatcatattcttattatgt-----caacaagaagaa 1081
 QY 791 ccacaattcagagagacgtgacgtgacgtcatatcctcatgtaattgagatg 850
 DB 1082 cacacacaacggaacaaaagctgagt-----gaaacttattactcgtttgat 1127
 QY 851 ctgtctgacagtgaagaagcacaagcgaactgagcgtgaagggtgtgagacttaagaagc 910
 DB 1128 ttagctcgtgtagtaaaaggttagtaaaactgagcgtgaagggtgtgtcgtgagtagc 1187
 QY 911 tgcacataaacccagcgtgtttatccttgagcaggttatataagaagccttaagcggcg 970
 DB 1188 aaaaataatacaaatcactctctgctctggaatgattatcttgcttgagtgaggg- 1246
 QY 971 cagcgtgtaggttataactacagacagacagaacacccagatctccaaatctta 1030
 DB 1247 -----tagtataatgttcacatacgagatagtaaaatgacaagaatcctcagaattca 1301
 QY 1031 ttggagagaatgtcaaacggttataattgtgacaataatgacagcttcttttgatgag 1090
 DB 1302 ttagtgagcaactgtagaaccactatgttaattgtgtcctccatcatcatacaatgag 1361
 QY 1091 acctctaa-----gtacacttcagttgcagcagtaacgtccaacaatgttgaaatctacc 1144

DB 1362 tctgaacaaaatctacacccctattctggccaaaggccaaacaattaagaacagct 1421
 QY 1145 catgttaataggttccctgtagatgaaagcgttctaaaaaggctacagaagaagaa 1199
 DB 1422 tgtgtcattgttgaggttaactctgcagacacgctggaaaaaagatctgaaaaagaaa 1476

RESULT 9
 AAS61626 standard; cDNA: 501 BP.

AC AAS61626:
 XX
 DT 29-JAN-2002 (first entry)
 XX

DE Lung small cell carcinoma antigen, cDNA #167.

KW Human; cytosolic; antitumour; lung small cell cancer antigen;
 KW tumour; lung cancer; ss.

OS Homo sapiens.
 XX

PN WO200177168-A2.
 XX

PD 18-OCT-2001.
 XX

PF 11-APR-2001; 2001WO-US11859.
 XX

PR 11-APR-2000; 2000US-196780P.
 XX

PR 21-JUN-2000; 2000US-213361P.
 XX

PR 01-SEP-2000; 2000US-229763P.
 XX

PR 05-SEP-2000; 2000US-230629P.
 XX

PR 14-SEP-2000; 2000US-232565P.
 XX

PR 19-DEC-2000; 2000US-257037P.
 XX

PR 08-JAN-2001; 2001US-260796P.
 XX

PA (CORI-) CORIXA CORP.
 XX

PI Lodes MJ, Wang T, Mohamath R, Indirias CY;
 XX

DR MPI; 2002-010896/01.
 XX

XX Lung tumour polynucleotide and polypeptides useful in therapy and
 PT diagnosis of cancer especially lung cancer .

PS Claim 1: Page 180; 295pp; English.

CC The invention relates to novel isolated lung small cell cancer antigen
 CC polynucleotides (I) and polypeptides (II) used in a method of detecting
 CC cancer in a patient. The method is optionally performed by
 CC utilising oligonucleotides (III), where the biological sample
 CC from the patient is contacted with (III), detecting the amount of
 CC polynucleotide hybridised to (III) in the sample and comparing the
 CC amount of polynucleotide to a predetermined cut-off value and thereby
 CC determining cancer in a patient. (I), (II) or antigen-presenting cells
 CC expressing (II) is useful for stimulating and/or expanding T cells
 CC specific for a tumour protein. The method comprises contacting T cells
 CC with one of the components under conditions to permit the stimulation
 CC and/or expansion of the cells. A composition comprising (I) is useful for
 CC stimulating an immune response in a patient and for inhibiting the
 CC development of a cancer especially lung cancer in a patient. An
 CC isolated T cell population is useful for removing tumour cells from the
 CC biological sample and for inhibiting the development of cancer in a
 CC patient. AAS61460-AAS61874 represent novel human lung small cell
 CC cancer antigen coding sequences of the invention.

Sequence 501 BP; 205 A; 67 C; 101 G; 128 T; 0 other;

Query Match 1.4%; Score 138.2; DB 24; Length 501;
 Best Local Similarity 57.7%; Pred. No. 1.5e-21;
 Matches 266; Conservative 0; Mismatches 193; Indels 2; Gaps 1;

QY 257 gttgatggacaaagtccttcaatttcgatcgtgtatttaattctcacgaatcaacaagt 316

[illegible]

QY 1137 atactcccatgttaatgaagtcctgagt 1166
 Db 1111 accacgctgtgtgaatgagaccctaag 1140

RESULT 13

AA22377
 ID AAF22377 standard; cDNA; 1701 BP.

XX AAF22377;

DT 26-MAR-2001 (first entry)

XX Human secreted protein gene 5 SEQ ID NO:15.

Human: secreted protein; diagnosis; immunosuppressive; antiarthritic;
 antiinflammatory; antiproliferative; cytostatic; cardiant; vasotropic;
 cerebroprotective; neurotropic; neuroprotective; antibacterial; vitruclide;
 fungicide; ophthalmological; gene therapy; autoimmune disease; neoplasm;
 rheumatoid arthritis; hyperproliferative disorder; cardiac arrest;
 cardiovascular disorder; cerebrovascular disorder; cerebral ischemia;
 anglogenesis; nervous system disorder; Alzheimer's disease; infection;
 ocular disorder; corneal infection; wound healing; skin aging;
 food additive; preservative; ss.

OS Homo sapiens.

PN WO200061629-A1.

PD 19-OCT-2000.

PF 06-APR-2000; 2000MO-US09071.

PR 09-APR-1999; 99US-0128694.

PR 20-JAN-2000; 2000US-0176931.

XX (HUMA-) HUMAN GENOME SCI INC.

PA (ROSE/) ROSEN C A.

PI Ruben SM, Komatsoulis G;

DR WPI: 2000-647420/62.

DR P-PSDB; AAB63138.

PT Isolated nucleic acid molecule encoding a human secreted protein is

PS Claim 1; Page 419-420; 533pp; English.

AA22377 to AAF22421 encode the human secreted proteins given in AAB63134
 CC to AAB63182. AAB63183 to AAB63231 represent more human secreted proteins
 CC and polypeptides homologous to them. Human secreted proteins have
 CC activities based on the tissues and cells the genes are expressed in.
 CC Examples of activities include: immunosuppressive; antiarthritic;
 CC antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic;
 CC cerebroprotective; neurotropic; neuroprotective; antibacterial; vitruclide;
 CC fungicide; and ophthalmological. The polynucleotides and proteins can be
 CC used to prevent, treat or ameliorate a medical condition in e.g. humans,
 CC mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They are
 CC also used in diagnosing a pathological condition or susceptibility to a
 CC pathological condition. Disorders which are diagnosed or treated include
 CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
 CC disorders e.g. neoplasms of the breast or liver, cardiovascular
 CC disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral
 CC ischemia, anglogenesis, nervous system disorders e.g. Alzheimer's
 CC disease, infections caused by bacteria, viruses and fungi and ocular
 CC disorders e.g. corneal infection. The polypeptides can also be used to
 CC aid wound healing and epithelial cell proliferation, to prevent skin
 CC aging due to sunburn, to maintain organs before transplantation, for
 CC supporting cell culture of primary tissues, to regenerate tissues and in
 CC chemotaxis. The polypeptides can also be used as a food additive or
 CC preservative to increase or decrease storage capabilities. AAF22364 to
 CC AAF22372 and AAB63133 represent sequences used in the exemplification of

CC the present invention.
 XX
 SQ Sequence 1701 BP; 516 A; 352 C; 437 G; 390 T; 6 other;

Query Match 1.3%; Score 123.2; DB 21; Length 1701;
 Best Local Similarity 50.0%; Pred. NO. 6.6e-18;
 Matches 435; Conservative 3; Mismatches 396; Indels 36; Gaps 4;

QY 321 ttacccaagaatagacagatcctacatcatalacagatcttgcagagatataatgaccacaa 380
 Db 275 ttctcaagtccttgagagagaatctctgcagagatctttrrattgctacatgcatgta 334
 QY 381 tattgcatacagacagacatcttcaggaagcgtacacaaatgaaatggaaccacaaatc 440
 Db 335 tcttgcctatgacagacatctgccttcgaaatcttatacatgaltgagcagatgacc 394
 QY 441 catggacataatcccccaagccatagagaggttttaaatattatcagggagataccga 500
 Db 395 aacctgatttaattcccaagacttgcagctggaaccttgcagacacacagaaagaggaa 454
 QY 501 acagagag-----ttcttctaagagttcttataatgagatltacaatgaaactgaa 554
 Db 455 atgagagacagagattttaaagtagaagtgcttccatcaggaatttataatgaaagagttc 514
 QY 555 aagacctactgtgtgatgacagagaagaagcccttggaatctgcgaggaatttataa 614
 Db 515 gaggaccttctgtatcccaagaagcgcgtcagacgtttgaaatgacagacatagtggt 574
 QY 615 gaaacgctgtgtgtgcagacacttgagaacttgcgttgaatggtcccggaacatgttaac 674
 Db 575 tgggaacctttagtcgacgagacttctaactgctgtccacagctacaagatgatatgagt 634
 QY 675 agtgcatacaaaagagtgtaaaagaaacagacatltatggagagacataaaatgaaatcata 734
 Db 635 cgtgatgctctgaggttaacaatctgcacagctgtgctgcaacaaatgaaagagagga 694
 QY 735 gtatgcttcaatataatlttgaatgattgttgaaagccgagacagaaatgattccca 794
 Db 695 gtagccgatacccatgacagtttcaaaataccacct-----cacacatactctc 742
 QY 795 caaatcagagaactgtgatgagagctgtcatgtatctcacttgaaatttgtagatcttg 854
 Db 743 acgattgtgaagctctggacatcttgagagaagaagtgggcaanstcagctgctgattag 802
 QY 855 ctggcagtgaaagagcaagcaaacctgagctgaaagtgtagaacttaagaaagctgca 914
 Db 803 mtggcagtgaaagagcaagcaagcagcgtctgcagagggagacaggtctaaagaaaggagca 862
 QY 915 acatcaaccgagctgtttatctcttgagacaggttataagaaagcttagcagcggcag 974
 Db 863 acattacaagctccctcacaaacctgctgtgtatcttcacgtctctgcagatcagagtg 922
 QY 975 ctgg-----tgattataactacacagagacacaaactcacagaattcttc 1022
 Db 923 ctggcaaaagcargaaataattgttccacatctgactcaagtctcactctgctgctca 982
 QY 1023 aaatctatggggagaaatgctaaacgggttaaatattgcaaaattgacgcagttctt- 1081
 Db 983 aagacagctctgggggaaacagcaagcagmccgcacatggtggtactctgtagtccgcagctg 1042
 QY 1082 -----ttgtatgagacttaagtaacttaagtttgcagtaactgcaacaaatgtagaa 1136
 Db 1043 ataactatgttgaacaccttcaacatctgtgtagatgagatcgagcaagcaactgttaa 1102
 QY 1137 atactcccatgttaatgaagtcctgagt 1166
 Db 1103 accacgctgtgtgaatgagaccctaag 1132

RESULT 14

AA22907
 ID AAC42907 standard; DNA; 3366 BP.

```
XX AAC42907;
AC
XX 17-OCT-2000 (first entry)
DT
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 37306.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KM metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
PN EPI033405-A2.
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128233.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136392.
PR 01-JUN-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137502.
PR 08-JUN-1999; 99US-0137724.
PR 10-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138547.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 27-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145813.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 06-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 09-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149475.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
```



```
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153700.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0158294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 23-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 1.2%; Score 117.6; DB 21; Length 3366;
Best Local Similarity 54.2%; Fred. No. 1.7e-16;
Matches 412; Conservative 0; Mismatches 309; Indels 39; Gaps 7;

```
OY 471 aagttttaaattatctcggagatccgaacagagattcttcttaagattcttata 530
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 830 aggaatgaatgcttcttaacggactccgaagagaggttctctgctgcttata 889
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 531 tggagattacaatgaaactgtgaaagactatgtgtatgacgaagaagaagccct 590
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 890 ttgaatatatacatgagtgatgaatgattactg-----gatccaacagccaaact 943
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 591 tggaaatccgagagatttataagaacgtgtatgtctgacccgtgactgagaactg 650
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 944 taagtgttagagagagattcccgagg--cacttatgttgaaggtatcaaggagaagt 1000
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 651 taatgttccctaacaatgaatacagtgatcaaaaaaggtgaaaaaacagacattat 710
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1001 ttgtgtccctcgccatgcactatcatctatgtcagctgagggaagaactgtcatgt 1060
```

```
OY 711 gagagactaaatgaatgatcatgagtgtgtccatatcaaatattagaatgtgtg 770
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1061 gtccaataatttcaattctgtgtgacgagagaagtcacacatatctaaacgtat 1120
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 771 aaagccgagacagaatgatccccaatctcaagaactgtgacggagctgacgtat 830
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1121 aaagtatgtg-----ctaacgtgagatgatatgtgtgaa---gtatctctt 1162
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 831 ctcaattgaatttgttagatctgtctgtgcaagtgaagaagcaagccaactgtgag 890
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1163 ctcaactcaatttgaattgactgtgacgtgtgaggttcga--aaactgaaacaact 1219
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 891 gtgtgagacttaagaagagctgtcaacatacaaccgagctgtgttatccctgtgac 950
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1220 gattgagaggaaggggttccatataccaagaaggtcttcaactctgtgaaactgt 1279
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 951 ttaagaagcttaagcagccgacgagctgtgtgattcttaactacagagaacaaactca 1010
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1280 ttggaaaacttaagtgaagggttaaggcaactcacattcca---tatcgtgactctaagt 1336
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 1011 ccagaattctccaaaattcaattgtggaagaatgtctaaacggtataatltgcaca 1070
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1337 ctctgtctctgcacatcttcaattgaatgtgtcatatgtgtcgtcatatgtaca 1396
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 1071 cggcaagttc-----tttgtatgactcttaagtcaacttgcactgttgcagactgtgca 1124
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1397 ctctctgtccgcagcagtaagtgaagaactcatalaacaactgtgaagttgtccaggaaggcaa 1456
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 1125 aaacatgtgaaatactcccatcgttaatgaggtccctgtgatgtgaaaggtgtcta 1184
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1457 agagtataagaataatatacgtcttcacgcaatcagaattataagatggaagatcaltat 1516
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 1185 ggtacagaagaagaaatcttgatttaagaagaacaattaga 1224
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1517 aatatcaaaagagaatctcaaccctcaaaactgaactgtga 1556
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
RESULT 15
AAH15830
ID AAH15830 standard; cDNA; 1950 BP.
XX
AC AAH15830;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:14319.
XX
KW Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99UP-0248036.
PR 27-AUG-1999; 99UP-0300253.
PR 11-JAN-2000; 2000UP-0118776.
PR 02-MAY-2000; 2000UP-0183767.
PR 09-JUN-2000; 2000UP-0241899.
XX
XX (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT primer sets for synthesizing polynucleotides, particularly the 5602
full-length cDNAs defined in the specification, and for the detection
```


THIS PAGE BLANK (USPTO)